सत्यमेव जयते

INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI

24959

I.A.R.I.6.
GIP NLK—H-3 I.A.R.I. 10 5 55—15,000

ANNALS

OF

THE NEW YORK ACADEMY OF SCIENCES

VOLUME XLIX

NEW YORK
PUBLISHED BY THE ACADEMY
1947–1948

Editor
ROY WALDO MINER

Associate Editor
LOTHAR SALIN

Consulting Editors
CHARLES GLEN KING
(Pages 1–140)

HAROLD G. CASSIDY
(Pages 141–326)

E. NEWTON HARVEY
(Pages 327–482)

GREGORY SHWARTZMAN
(Pages 483–660)

ROBERTS RUGH
(Pages 661–866)

S. STANSFELD SARGENT
(Pages 867–928)

CONTENTS OF VOLUME XLIX

PAGE

Title Page i

Contents iii

Nutrition in Relation to Cancer. *By* C. G. KING, J. J. BITTNER, D. H. COPELAND, R. W. ENGEL, C. J. KENSLER, G. W. KIDDER, G. B. MIDER, J. A. MILLER, H. P. MORRIS, W. J. ROBBINS, W. D. SALMON, A. TANNENBAUM, E. L. TATUM, F. R. WHITE, J. WHITE, AND P. R. WHITE 1

Chromatography. *By* H. G. CASSIDY, N. APPLEZWEIG, S. CLAESSON, V. R. DEITZ, B. J. MAIR, A. J. P. MARTIN, S. MOORE, R. L. PECK, W. A. SCHROEDER, L. SHEDLOVSKY, W. H. STEIN, H. C. THOMAS, AND L. ZECHMEISTER 141

Bioluminescence. *By* E. N. HARVEY, R. S. ANDERSON, J. B. BUCK, A. M. CHASE, H. EYRING, AND F. H. JOHNSON 327

Hemorrhage. *By* G. SHWARTZMAN, C. H. BEST, R. CHAMBERS, C. S. DAVIDSON, J. H. FERGUSON, F. F. FURCHGOTT, I. E. GERBER, M. I. GREGERSEN, P. GYÖRGY, R. L. HADEN, C. HYMAN, L. B. JAQUES, P. KLEMPERER, R. E. LEE, C. N. H. LONG, D. W. RICHARDS, JR., R. H. SCHNEIDER, E. SHORR, H. J. TAGNON, S. A. THAYER, L. C. UNDERWOOD, D. D. VAN SLYKE, A. E. WILHELMI, AND B. W. ZWEIFACH 483

Recent Studies in the Mechanisms of Embryonic Development. *By* E. J. BOELL, D. P. COSTELLO, S. R. DETWILER, G. FANKHAUSER, J. HOLTFRETER, J. S. NICHOLAS, S. M. ROSE, D. RUDNICK, AND L. S. STONE 661

Current Trends in Clinical Psychology. *By* A. W. COMBS, H. E. DURKIN, M. L. HUTT, J. G. MILLER, J. L. MORENO, AND F. C. THORNE 867

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
VOLUME XLIX, ART. 1. Pages 1-140
SEPTEMBER 7, 1947

NUTRITION IN RELATION TO CANCER *

By

CHARLES GLEN KING, JOHN J. BITTNER, D. H. COPELAND, R. W. ENGEL, C. J. KENSLER, GEORGE W. KIDDER, G. BURROUGHS MIDER, J. A. MILLER, HAROLD P. MORRIS, WILLIAM J. ROBBINS, W. D. SALMON, ALBERT TANNENBAUM, E. L. TATUM, FLORENCE R. WHITE, JULIUS WHITE, and PHILIP R. WHITE

CONTENTS

PAGE
Introduction to the Conference. *By* CHARLES GLEN KING 3
Effects of Varying Caloric Intake upon Tumor Incidence and Tumor Growth. *By* ALBERT TANNENBAUM 5
Studies on the Mechanism of the Effects of Fats and other Dietary Factors on Carcinogenesis by the Azo Dyes. *By* J. A. MILLER 19
Effect of Diet on the Production of Liver Tumors in the Rat by N,N-Dimethyl-p-Aminoazobenzene. *By* C. J. KENSLER 29
Level of Protein Ingestion and an Appraisal in Terms of Protein Composition. *By* JULIUS WHITE, FLORENCE R. WHITE, and G. BURROUGHS MIDER 41
Carcinogenic Effects Associated with Diets Deficient in Choline and Related Nutrients. *By* R. W. ENGEL, D. H. COPELAND, and W. D. SALMON... 49
The Mammary Tumor Milk Agent. *By* JOHN J. BITTNER 69
Some Basic Contributions to the Cancer Problem from the Study of Plants. *By* WILLIAM J. ROBBINS 75
Chemically Induced Mutations and their Bearing on Carcinogenesis. *By* E. L. TATUM 87
The Nutrition of Monocellular Animal Organisms. *By* GEORGE W. KIDDER 99
The Nutrition of Malignant Tissue *in vitro*. *By* PHILIP R. WHITE 111
Effects on the Genesis and Growth of Tumors Associated with Vitamin Intake. *By* HAROLD P. MORRIS 119

---

* This series of papers is the result of a Conference on Nutrition in Relation to Cancer held by The New York Academy of Sciences and the Panel on Nutrition, Committee on Growth, National Research Council, on December 6 and 7, 1946.

Publication aided by a grant from The Nutrition Foundation, Inc., New York, N. Y.

COPYRIGHT 1947

BY

THE NEW YORK ACADEMY OF SCIENCES

INTRODUCTION TO THE CONFERENCE

By Charles Glen King

The Nutrition Foundation, Inc., New York, N. Y.

Research progress in the study of cancer is terribly discouraging to many people. This is very understandable because, aside from surgery and radiation therapy, so little of practical value or in the direction of basic understanding of the problem seems to have been accomplished. Yet we can scarcely doubt the successful outcome of intensive research effort. The limitation is primarily one of time and manpower. The challenge in cancer research, as measured in human values and in complexity of attack, is surely as great as in any field that a scientist could find.

Basically, we know that attempts to use shortcuts are likely to lead to disappointments, but we also know that there is need to have the viewpoint of both the meticulous plodder and the impatient driver, *i.e.*, those who have been aptly characterized as "single variabilists" and "multiple variabilists".

It may seem disappointing that, after so much study of cancer, no one knows, in a fundamental way, the cause or nature of tumor development. Accordingly, a rational approach to prevention or cure seems extremely remote. Yet I think we should be greatly encouraged, especially in nutrition research, by two findings that are firmly established, since they provide powerful tools with which to work:

(1) A variety of definite, pure compounds are known that apparently initiate carcinogenesis, so that one can lay out a research program on a molecular basis. Unfortunately, we do not know whether these compounds are similar to the common causative agents in human cancer.

(2) Under controlled conditions with animals, several substances that are characteristic of natural diets can completely decide the issue of whether or not tumors will develop. Recent results with choline in rat feeding illustrate this in a striking manner.

Besides the two major leads just cited, I think we have reason to be encouraged by the discoveries that have been made in regard to caloric intake, percentage and composition of ingested fat, certain members of the vitamin B-complex (especially riboflavin), the more complex agents that can be transmitted in milk, and the availability of tracer elements with which to explore what happens inside normal and cancer cells.

The New York Academy of Sciences and the Committee on Growth of the National Research Council are greatly indebted to those who have prepared papers for the program and to those who have consented to preside at the three main sessions. The research men who are working on widely scattered approaches to the cancer problem will, I am sure, enjoy both the formal papers and the discussions that will follow.

Before turning to their presentation, I want to express a personal word of thanks to Mrs. Miner, Doctor Owen, Doctor Rhoads, Doctor Nigrelli, and to members of the Panel on Nutrition, Doctor Elvehjem, Doctor Woolley, and Doctor Vickery for their assistance in organizing the program.

EFFECTS OF VARYING CALORIC INTAKE UPON TUMOR INCIDENCE AND TUMOR GROWTH

By ALBERT TANNENBAUM

Department of Cancer Research,
Michael Reese Hospital, Chicago, Illinois

That tumors might be dependent, in part, on nutrition, has been surmised for many years. This viewpoint was based, no doubt, on the simple but significant fact that nutrition is needed for the origin and growth of both individual cells and organisms. At first, the approach was speculative; in the past few years, however, the experimental attack has furnished much valuable information.

The relationship of nutrition to cancer has many facets. These are discussed in the original publications and in a review presented at the Gibson Island Cancer Conference of 1945 (Tannenbaum, 1945c). In fact, this limited presentation on the effects of varying the caloric intake upon tumor incidence and tumor growth has been taken, mainly, from the latter review.

The origin or genesis of a tumor must be differentiated from its subsequent growth. Actually, these are two phases of the tumor process and they may be affected in different ways by an experimental procedure or agent. It is for this reason that incidence and growth of tumors are discussed separately.

EFFECTS OF CALORIC INTAKE ON THE INCIDENCE OF TUMORS

The fact, now established, that caloric intake is a factor in the production of tumors was suggested to us in experiments, started in 1937, designed to evaluate the effects of various common food constituents. Although all the animals in these experiments were being fed *ad libitum*, some weighed less than others, and it appeared that fewer of the smaller animals developed tumors. These smaller mice had free access to food, but the exact food consumption in comparison with that of the others in the groups was not known. It was for these reasons that experiments were initiated to test the effect of simple underfeeding (giving the experimental mice about two-thirds of the ration consumed by the controls). It was found that simple underfeeding inhibited the formation of tumors, and the evidence suggested that the effect was produced by restriction of calories rather than by some specific food component. Therefore, other experiments were begun in which the restricted diet differed from the *ad libitum* diet in caloric (carbohydrate) value only.

In PLATE 1, 10 mice of a typical experiment are seen. Five are from the control group and five from the underfed group. The underfed mice appeared normal, clinically, except that they were smaller and more active, particularly before feeding time, than the control mice. Actually, the underfed mice lived longer, on the average. Another interesting feature was that the underfed mice revealed fewer pathologic changes of the heart, kidneys, liver, etc., than did the control mice of the same age.

What are the actual effects of simple underfeeding or caloric restriction on the incidence of tumors? A few typical experiments will be detailed, following which the main work in this field will be summarized. TABLE 1 illustrates the striking effect on the formation of spontaneous mammary carcinoma in DBA virgin female mice. The mice of the N45 restricted group were fed and ate 2.0 grams of the basic ration, their average weight being maintained at from 19 to 21 grams throughout the course of the experiment. On the other hand, the mice of the N42 control group were given the same amount of basic ration plus corn starch. They ate, on the average, 3.0 grams of food daily, and grew at a normal rate.

TABLE 1

THE EFFECT OF A CALORIE-RESTRICTED DIET ON THE FORMATION OF SPONTANEOUS MAMMARY TUMORS IN DBA VIRGIN FEMALE MICE

	N42: *ad-libitum* controls			N45: caloric-restricted		
Age weeks	Mean weight gm.	Animals alive and tumor-free	Cumulative tumor count	Mean weight gm.	Animals alive and tumor-free	Cumulative tumor count
10*	20	50	0	21	50	0
40	28	48	0	19	50	0
48	30	47	1	19	50	0
56	30	45	2	20	49	0
64	31	40	6	21	48	0
72	30	35	11	20	46	0
80	29	27	16	20	41	0
88	—	15	23	—	39	0
96	—	8	25	—	31	0
100	—	6	26	—	29	0

* Animals placed on experimental diets at 10 weeks of age.

By 100 weeks of age, 26 animals in the group fed *ad libitum* had developed tumors, while none had appeared in the restricted group. This striking effect was not caused by deaths at an early age in the restricted group, for it is seen that, at 1 year, there are actually 49, and at 100 weeks there are still 29, of the original 50 mice. These results clearly show that a caloric-restricted diet inhibited the formation of spontaneous mammary carcinoma in the mouse.

TABLE 2

THE EFFECT OF A CALORIE-RESTRICTED DIET ON THE FORMATION OF INDUCED SKIN TUMORS IN DBA MALE MICE*

	A0: *ad-libitum* controls			A5: caloric-restricted		
Weeks after first application*	Mean weight gm.	Animals alive and tumor-free	Cumulative tumor count	Mean weight gm.	Animals alive and tumor-free	Cumulative tumor count
†	23	50	0	22	50	0
0	27	50	0	20	50	0
8	31	49	0	22	50	0
16	34	41	8	21	50	0
24	33	34	15	21	47	2
32	33	26	21	21	39	5
40	32	19	25	22	37	6
48	32	10	29	22	33	9
56	—	4	32	21	29	11
60	—	4	32	20	26	11

* Nineteen semi-weekly applications of benzpyrene.

† Diets started 4 weeks before first application of carcinogen.

In TABLE 2, the effect of a caloric-restricted diet on the formation of induced skin tumors is shown. The mice of both groups received the same amount of carcinogen, 19 semi-weekly applications of a 0.3 per cent solution of 3:4-benzpyrene in benzene given by means of a dropper to the skin of the interscapular area. The experiment was continued for 60 weeks after the first application of the carcinogen. The mice of the restricted group (A5) received 2.3 grams of the basic ration, while those of the control group (A0) received the same amount of basic ration plus 1.9 grams of corn starch. The mice of the restricted group maintained an average weight of 20 to 22 grams, while those of the *ad libitum* control group grew normally, reaching a maximum average weight of 34 grams. Only 11 tumors arose in the restricted group, in comparison with 32 in the control group fed *ad libitum*. Thus, a caloric-restricted diet inhibited the formation of skin tumors induced by a carcinogenic hydrocarbon.

FIGURE 1 represents the results obtained in an experiment with the sarcoma induced by 3:4-benzpyrene (a single subcutaneous injection, in the interscapular area, of 0.2 mg. of the carcinogen in 0.2 cc. of an oily fraction of lard). The curves of cumulative number of mice with induced sarcomas are given for the control group (N12) and the corresponding calorie-restricted group (N15). Here, again, the inhibiting effect of a calorie-restricted diet is seen.

The results of these three typical experiments, utilizing different types of tumors, are representative of many experiments performed in this laboratory. In every investigation, it was found that caloric restriction

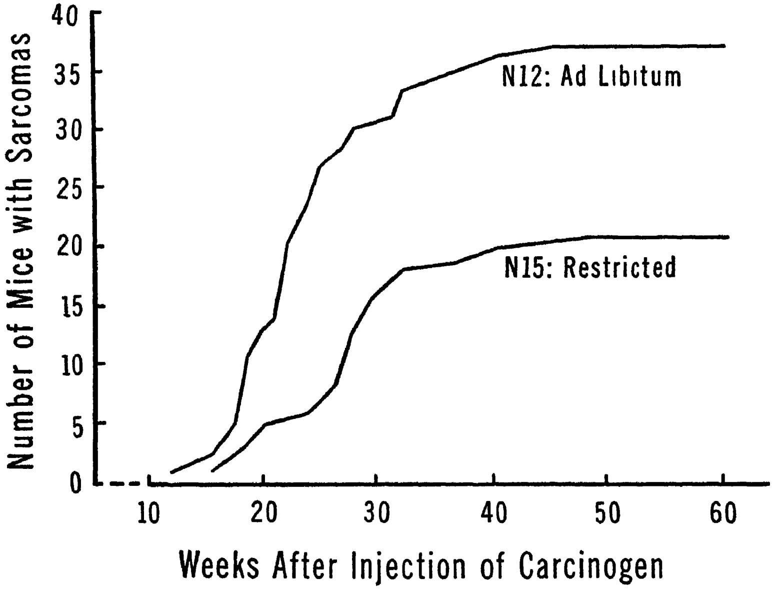


FIGURE 1. Inhibition of the formation of induced sarcomas by means of a calorie-restricted diet. 50 animals in each group at beginning of experiment.

inhibited the formation of tumors. This inhibition was characterized by both *a decrease in the total number of tumors* and *a delay in the average time of appearance of the tumors.*

FIGURE 2 is a general summary chart of the investigations concerned with the inhibitory effect of underfeeding or caloric restriction on tumor genesis. It represents the experiments performed in our laboratory as well as those of other investigators. The chart was prepared before the publication of the experiments of Larsen and Heston on spontaneous pulmonary tumors (1945), those of Rusch, Johnson, and Kline (1945b) on induced sarcomas, and our experiments with the spontaneous hepatomas in male C3H mice. The results of these investigations are not given in the figure, but are in agreement with those shown.

The chart depicts the effects of underfeeding or caloric restriction on five different types of tumors in mice: spontaneous mammary carcinoma, induced skin tumor, induced sarcoma, primary lung adenoma, and leukemia. The total height of the bar (cross-hatched and black) represents the incidence of tumors in the *ad libitum* or full-fed control group, while the lower component of the bar (cross-hatched) represents the incidence of tumors in the experimentally restricted group. The black portion of the bar, therefore, represents the inhibiting effect of the restricted diet.

Those bars which have an experiment number below represent investigations performed in this laboratory; those of other investigators are indicated by an appropriate initial. When there is no letter above the bar, caloric restriction was achieved by limiting all components proportionately (simple underfeeding), by giving the restricted mice a fraction of the *ad libitum* ration. When C is above the bar, it implies carbohydrate restriction only; C+F above the bar indicates that the restriction was achieved by limiting both carbohydrate and fats.

It is to be noted that all five types of tumors respond to caloric restriction in the same way: fewer mice develop tumors. In the early experiments (Tannenbaum, 1940a, 1942) we chose four different kinds of tumors. These randomly selected types of tumors, readily available for investigation, were all affected by caloric restriction. It appeared, therefore, that there must be some generality in the effect. Since then, excellent investigations by Visscher and associates (1942), Saxton, Boon, and Furth (1944), White and her associates (1944), Rusch and associates (1945a, 1945b), and Larsen and Heston (1945) have confirmed these observations and have extended them to other tumor types and to spon-

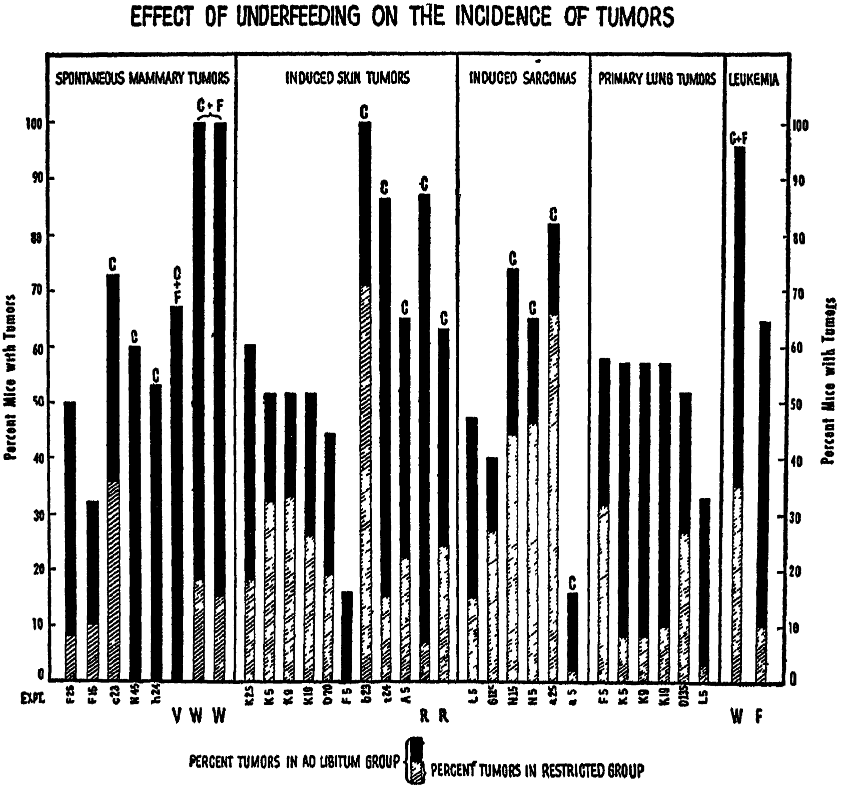


FIGURE 2. Effect of restricted food intake on the incidence of various types of tumors in mice. See text for general description of chart. F: Saxton, Boon, & Furth, 1944; R: Rusch et al., 1945a, V: Visscher et al., 1942, W: White et al., 1944. The remaining columns were drawn from Tannenbaum, 1940a, 1942, 1945a, 1945b; experiments N5, a25, a5 from unpublished data.

taneous and induced leukemia. We now have data indicating that the spontaneous hepatoma of the C3H male mouse is affected in the same way (Tannenbaum, 1945b). Thus, at least 8 different types of tumors and leukemia of the mouse are inhibited by caloric restriction. Furthermore, as yet no tumor has been found that does not respond in this way.

There are probably many factors that affect the actual degree of inhibition that occurs in a particular experiment. For example, it is likely that the kind of tumor is such a factor (see FIGURE 2). Without discussing the details of each experiment, it appears that, as a whole, the spontaneous mammary and lung tumors, and spontaneous hepatoma are inhibited by caloric restriction to a greater extent than are the induced skin tumors which, in turn, respond more readily than the induced sarcoma. These differences may be due to the relative malignancy of the tumors. The dosage of carcinogen (endogenous or exogenous), the degree of caloric restriction, to be discussed next, and other experimental conditions undoubtedly are important in the modification of the extent of the caloric effect. With the proper choice of tumor or dose of carcinogen, and a greatly restricted caloric intake, tumor formation may be inhibited to the extent that none appears. However, it is more important to stress the general nature and effect of caloric restriction than to emphasize the sensational effect (no tumors at all) that may occur through a combination of proper experimental conditions and drastic caloric restriction.

Effect of Varying the Degree of Caloric Restriction. Most of the investigations on the effect of caloric restriction have been performed by feeding the experimental group approximately 60 per cent of the caloric intake of the control group. It would be important if smaller degrees of restriction would also result in a significant effect. It was with this in mind that investigations utilizing graded caloric intakes were performed (Tannenbaum, 1945a, 1945b).

These experiments employed the spontaneous mammary tumor, spontaneous hepatoma (unpublished), and the induced skin tumor. The restricted diets, which differed from one another and from the *ad libitum* diet only in the amount of carbohydrate, ranged from approximately 60 to 90 per cent of the *ad libitum* diet, in caloric value. This was accomplished by feeding the restricted groups and the *ad libitum* group of a particular series the same amount of Purina Fox Chow meal and skimmed milk powder, and adding different amounts of corn starch to the several diets.

Though space does not permit a detailed review of these experiments, it can be said that, in general, the results indicate that the inhibition of tumor formation is dependent on the degree of caloric restriction, that is, the lower the caloric intake the greater the tumor inhibition. A tentative curve, based on the actual results, of the relationship of graded caloric

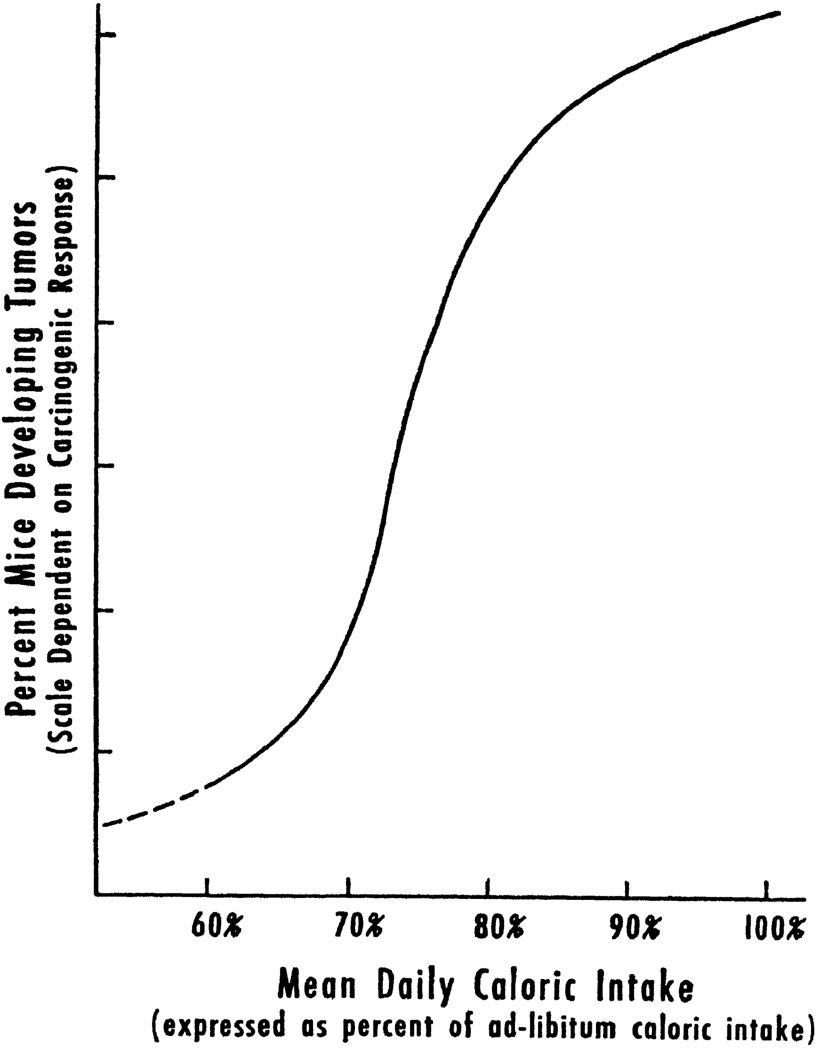


FIGURE 3. Tentative composite curve showing dependence of tumor incidence on degree of caloric restriction (restriction in carbohydrate only).

restriction to the formation of spontaneous mammary tumor, spontaneous hepatoma, and induced skin tumor is given in FIGURE 3. At the present time, the evidence suggests that the curve has a modified ʃ shape. Under certain experimental conditions, this may be quite obvious, while under others it may be difficult to distinguish from a linear relationship, particularly if only a few scattered levels of caloric restriction are studied.

Effect of Dosage of Carcinogen. The question arose as to whether the effect of caloric restriction on tumor formation might be obliterated in such instances where the carcinogenic stimulus is massive rather than moderate or mild. This has been tested in an experiment (Tannenbaum and Silverstone, 1946a) where 4 groups of mice were utilized: 2 of the groups were placed on a caloric-restricted diet, one receiving a moderate and the other a massive dosage of carcinogen. Each of these groups had its own control fed an *ad libitum* diet. The results are shown in TABLE 3. The effect of caloric restriction on tumor incidence observed with the moderate dosage of carcinogen was obliterated by the massive dosage. In contrast, the massive dosage actually emphasized the statistical validity of the difference in the mean time of appearance of the tumors. Thus, even with a massive dosage of carcinogen, the effect of caloric restriction on tumor formation is still apparent.

TABLE 3

THE EFFECTS OF CALORIC RESTRICTION AS INFLUENCED BY DOSAGE OF CARCINOGEN

0.3% Benzpyrene in benzene	Diet	Number of mice	Percentage of tumors	Time of appearance of skin tumors in weeks				
				Mean	Variance	Difference between means	Significance* of difference t	P
Moderate dosage (19 applications)	*ad-libitum* (A0)	49	65	28.3 ± 2.4	178	9.7 ± 4.6	2.1	<.05
	calorie-restricted (A5)	50	22	38.0 ± 4.0	172			
High dosage (50 applications)	*ad-libitum* (A60)	49	96	16.3 ± 0.5	10.3	4.4 ± 0.81	5.5	<.001
	calorie-restricted (A65)	48	90	20.7 ± 0.7	18.3			

* t and P refer to Fisher's test for the significance of the difference between two means (Fisher, 1938). The values for the groups given the massive dosage of carcinogen have "greater" statistical significance.

Caloric Restriction During Different Stages of Carcinogenesis. On three occasions, experiments had been performed on the effect of caloric restriction on the formation of mammary carcinoma in DBA females. The experiments differed in the type of diets used, but in each case a striking inhibition of tumor formation was observed. The data are briefly given in TABLE 4.

TABLE 4

EFFECT OF INSTITUTION OF RESTRICTED DIETS AT DIFFERENT AGES ON THE FORMATION OF SPONTANEOUS MAMMARY TUMORS IN DBA MICE

Experimental groups		Number of mice in group		Experimental diets started at (average age) months	Mice developing mammary tumors by 20 months of age	
Restricted	*ad libitum*	Restricted	*ad libitum*		Restricted per cent	*ad libitum* per cent
F15	F12	44	44	9	5	25
F25	F32	50	50	5	2	40
N45	N42	50	50	2	0	38

One noteworthy difference in the experiments was the age of the mice at the time the investigations (and the experimental diets) were started. This was 9, 5, and 2 months, respectively. The comparable inhibitory effects on tumor formation obtained in these 3 experiments suggested that the formation of spontaneous mammary tumors in the DBA strain can be inhibited if caloric restriction is instituted at any time before 9 months, which is the time when tumors begin to appear. This further suggested that the inhibitory effect of caloric restriction may be exerted chiefly during the period in which tumors appear, rather than in the previous period of carcinogenic preparation of the mammary tissue. This concept was confirmed in an experiment (Tannenbaum, 1944) in which different groups were all given carcinogen applications to the skin but varied as to the time at which the calorie-restricted diets were instituted. Caloric restriction produced its effect even when instituted just before the tumors began to appear.

Implications of Caloric Restriction. The inhibitory effect of caloric restriction on tumor formation has significance in other fields of cancer research. It is possible that a particular experimental procedure may retard the growth or decrease the food intake of the experimental animal. In such instances, any inhibitory effect on the formation of tumors may have no direct dependence on the experimental procedure, except as the latter affects body growth and caloric intake. For example, we have found that such diverse procedures as the incorporation of sodium fluoride or dinitrophenol in the diet, or keeping the experimental mice in a cold room (at about 50° F.) on a caloric intake equal to that of control mice, markedly inhibited the formation of spontaneous mammary tumors (Tannenbaum and Silverstone, 1946b). The sodium fluoride caused a voluntary restriction of food intake, while the other two procedures accelerated the metabolic rate (however, since food intake was not permitted to increase, the resultant effect was an imposed "restriction"). In all three cases, the animals weighed decidedly less than the control mice. It is very likely that the decreased incidence of mammary carcinoma was not principally dependent on the sodium fluoride, dinitrophenol, or the low environmental temperature, but was rather mediated through their effects on food intake, food requirement, and body growth.

EFFECTS OF CALORIC RESTRICTION ON THE GROWTH OF TUMORS

There are many reports on the retarding effect of underfeeding on the "growth" of transplanted tumors (Moreshi, 1909; Rous, 1914; Sugiura and Benedict, 1926; Bischoff and associates, 1935). Later, Bischoff and Long (1938) demonstrated that the inhibition of growth was due to caloric restriction *per se* (restriction of either starch or fat only). In at least some of these investigations, there was not a clear separation between the establishment and the growth of the tumors.

In this laboratory, the effects of either simple underfeeding or caloric restriction on the growth of tumors, already established, have been studied. Spontaneous mammary carcinoma arising in underfed or caloric-restricted mice appears to grow at approximately the same rate as tumors arising in full-fed controls (Tannenbaum, 1940a). On the other hand, the rate of growth of this same type of tumor arising in full-fed animals is significantly diminished when the animals are subsequently underfed.

Sarcomas induced by subcutaneous injection of 3:4-benzpyrene grow at only a slightly lower rate when formed in underfed animals (Tannenbaum, 1942). Sarcomas arising in full-fed mice grow at a slower rate after the animals are placed on a caloric-restricted diet, but the mean life span of the sarcoma-bearing mice is not increased by the underfeeding (unpublished).

In experiments with transmitted mouse leukemia, Flory, Furth, Saxton, and Reiner (1943) found that, with some strains of leukemia, underfeeding, as compared with an *ad libitum* diet, prolonged the lives of animals previously given intravenous inoculations of leukemia cells. In one experiment with myeloid leukemia 106, the average survival of underfed mice was 29 days, in comparison with 16 days for the control mice. On the other hand, in experiments with myeloid leukemia 1712, the average life span was unaffected or even shortened somewhat by underfeeding. In these investigations, the underfeeding was begun the day after inoculation, and the question arises as to whether the underfeeding delayed the establishment ("take") of the leukemia or actually inhibited subsequent proliferation (growth).

It appears that underfeeding or caloric restriction is not a *practical* means of affecting the growth of established tumors. Drastic caloric restriction can retard the growth of a tumor, but at the same time the body weight of the host also diminishes. It is questionable if the life span of the tumor-bearing animal can be significantly increased by this procedure.

POSSIBLE ROLE OF CALORIC RESTRICTION IN HUMAN CANCER

For many years, insurance companies have been studying the relationship of various factors to the principal causes of death. Some of these studies consider the relationship of body weight to cancer mortality. A typical study is that of Dublin, who used approximately 192,000 records (1887–1921) of the Union Central Life Insurance Company for an analysis of cancer mortality. In this study, the policy holders, men who had bought insurance at 45 years of age and over, were classified according to weight at issue of policy. The distribution of cancer mortality with regard to weight is shown in TABLE 5. This study reveals a sufficiently consistent gradation in mortality rates to indicate that cancer incidence increases with increasing weight.

TABLE 5

Weight at issue of policy	Cancer mortality per hundred thousand
25 per cent or more overweight	143
15 to 25 per cent overweight	138
5 to 15 per cent overweight	121
Normal weight	111
5 to 15 per cent underweight	114
15 to 50 per cent underweight	95

A review (Tannenbaum, 1940b) of this and five other available insurance statistics studies indicated that individuals who are overweight when past middle age are more likely to die of cancer than are persons of average weight or less. Coupling these results with those obtained with mice, it was concluded that the avoidance of overweight might result in the *prevention* of a considerable number of cancers in humans or, at least, the cancer process may be delayed in time of appearance.

It is to be emphasized that the foregoing statement refers to the formation of tumors. Present evidence suggests that it is unlikely that caloric-restriction is a practical means of affecting the growth of an established tumor. However, there are reasons for believing that caloric restriction might have an inhibiting effect upon the establishment of tumor emboli or, possibly, very recently established metastases.

SUMMARY

There is considerable experimental evidence that caloric restriction inhibits the genesis or incidence of all types of mouse tumors that have been studied, namely, induced skin tumors, induced sarcomas, spontaneous mammary carcinomas, spontaneous lung adenomas, spontaneous hepatomas, and spontaneous and induced leukemia. Fewer mice develop tumors, and these appear, on the average, at a later time. The inhibitory effect is dependent on the degree of caloric restriction, the type of tumor, and the dosage or potency of the carcinogen. It is probable that the main influence occurs during the development of the tumor rather than during the preparatory stage.

The growth of tumors can be inhibited by caloric restriction, but the host also loses weight. Present evidence does not suggest that caloric restriction may affect the growth of tumors in a practical, useful way.

BIBLIOGRAPHY

Bischoff, F., M. Louisa Long, & L. C. Maxwell
1935. Influence of caloric intake upon the growth of sarcoma 180. Am. J. Cancer **24**: 549.
Bischoff, F., & M. Louisa Long
1938. The influence of calories *per se* upon the growth of sarcoma 180. Am. J. Cancer **32**: 418.
Fisher, R. A.
1938. Statistical Methods for Research Workers. Oliver & Boyd. London.
Flory, C. M., J. Furth, J. A. Saxton, Jr., & L. Reiner
1943. Chemotherapeutic studies on mouse leukemia. Cancer Research **3**: 729.
Larsen, C. D., & W. E. Heston
1945. Effects of cystine and caloric restriction on the incidence of spontaneous pulmonary tumors in strain A mice. J. Nat. Cancer Inst. **6**: 31.
Moreschi, C.
1909. Beziehungen zwischen Ernährung und Tumorwachstum. Z. Immunitätsforsch. **2**: 651.
Rous, P.
1914. The influence of diet on transplanted and spontaneous mouse tumors. J. Exp. Med. **20**: 431.
Rusch, H. P., B. E. Kline, & C. A. Baumann
1945a. The influence of caloric restriction and of dietary fat on tumor formation with ultraviolet radiation. Cancer Research **5**: 431.
Rusch, H. P., R. O. Johnson, & B. E. Kline
1945b. The relationship of caloric intake and of blood sugar to sarcogenesis in mice. Cancer Research **5**: 705.
Saxton, J. A., Jr., M. C. Boon, & J. Furth
1944. Observations on the inhibition of development of spontaneous leukemia in mice by underfeeding. Cancer Research **4**: 401.
Sugiura, K., & S. R. Benedict
1926. The influence of insufficient diets upon tumor recurrence and growth in rats and mice. J. Cancer Research **10**: 309.
Tannenbaum, A.
1940a. The initiation and growth of tumors. Introduction. I. effects of underfeeding. Am. J. Cancer **38**: 335.
1940b. Relationship of body weight to cancer incidence. Arch. Path. **30**: 509.

1942. The genesis and growth of tumors. II. Effects of caloric restriction *per se*. Cancer Research **2**: 460.

1944. The dependence of the genesis of induced skin tumors on the caloric intake during different stages of carcinogenesis. Cancer Research **4**: 673.

1945a. The dependence of tumor formation on the degree of caloric restriction. Cancer Research **5**: 609.

1945b. The dependence of tumor formation on the composition of the calorie-restricted diet as well as on the degree of restriction. Cancer Research **5**: 616.

1945c. The role of nutrition in the origin and growth of tumors. A. A. A. S.-Gibson Island Cancer Conference. In: Approaches to Tumor Therapy. A. A. A. S., Washington, D. C.

Tannenbaum, A., & H. Silverstone

1946a. Significance of dosage of carcinogen in evaluating experimental procedures. Abstract. Cancer Research **6**: 501.

1946b. Effect of sodium fluoride, dinitrophenol, and low environmental temperature on the formation of spontaneous mammary carcinoma in mice. Abstract. Cancer Research **6**: 409.

Visscher, M. B., Z. B. Ball, R. H. Barnes, & I. Sivertsen

1942. The influence of caloric restriction upon the incidence of spontaneous mammary carcinoma in mice. Surgery **11**: 48.

White, Florence R., J. White, G. B. Mider, Margaret G. Kelly, & W. E. Heston

1944. Effect of caloric restriction on mammary tumor formation in strain C3H mice and on the response of strain DBA to painting with methylcholanthrene. J. Nat. Cancer Inst. **5**: 43.

PLATE 1

Five representative underfed mice (left) in comparison with five mice fed *ad libitum* (right).

TANNENBAUM: EFFECTS OF CALORIC INTAKE ON TUMORS

STUDIES ON THE MECHANISM OF THE EFFECTS OF FATS AND OTHER DIETARY FACTORS ON CARCINOGENESIS BY THE AZO DYES *

By J. A. MILLER

McArdle Memorial Laboratory, Medical School, University of Wisconsin, Madison, Wisconsin

One phase of the studies[1] by the Wisconsin group† on the carcinogenic azo dyes‡ has concerned the great effect that diet can exert on the activity of these substances for the liver of the rat. In our initial studies,[2, 3] we confirmed both the report of Nakahara *et al.*[4] that the inclusion of 10 per cent whole liver powder in the diet would greatly retard the development of liver tumors in rats fed p-dimethylaminoazobenzene, and the duplication of this effect in terms of high riboflavin and casein supplementation as later reported by Kensler *et al.*[5] Our subsequent work was directed towards (1) the development of a satisfactory purified diet for the production of liver tumors in the rat by feeding this azo dye and (2) an assay procedure for evaluating the effects produced by changes in this diet. For assay purposes, we have routinely fed the test diets containing 0.06 per cent of p-dimethylaminoazobenzene for 4 to 4½ months to young adult rats of the Sprague-Dawley strain,§ ascertained the tumor incidence and the degree of cirrhosis at this time by performing laparotomies, and made the final tumor count after 2 additional months on the same diet without the dye to allow latent tumors to develop. The composition of the purified diet which has proved satisfactory under these conditions is given in TABLE 1. In general, when 2 mg. of riboflavin were added per kg. of diet, a tumor incidence of about 30 per cent was obtained at 4 to 4½ months, although occasionally no tumors were detected at this time. By 6 to 6½ months, 50 to 80 per cent of the rats had developed neoplasms in the liver, which establishes a convenient range for the detection of both the stimulation and the inhibition of tumor development by changes in this diet. While most of our work has been done with dye-feeding periods of 4 months followed by 2 months on the dye-free diets, within the past year and a half the Sprague-Dawley rat has appeared to be somewhat less susceptible to tumor development with this diet. Accordingly, we now either feed the diet containing 2 mg. of

---

* These studies were aided by grants from the Jane Coffin Childs Memorial Fund for Medical Research, and from the National Cancer Institute.

† Most of this research has been carried out by Dr. E. C. Miller, Mr. B. E. Kline, Prof. H. P. Rusch and myself at the McArdle Memorial Laboratory, and by Prof. C. A. Baumann and his associates in the Department of Biochemistry.

‡ For a summary of the research on these dyes through 1944, see reference [1].

§ Sprague-Dawley, Inc., Madison, Wisconsin.

riboflavin per kg. for an extra 2 weeks, or lower the riboflavin content to about 1.5 mg. per kg. to obtain the tumor incidences previously reported.

TABLE 1

CONTROL DIET FOR MEDIUM TUMOR INCIDENCE
(grams per kilogram)

Casein (vitamin low)	120
Glucose	790
Corn oil	50
Salts mixture	40
p–Dimethylaminoazobenzene	0.6
Riboflavin	0.001-0.002
Thiamine hydrochloride	0.003
Pyridoxine hydrochloride	0.0025
Calcium pantothenate	0.007
Choline chloride	0.030
Halibut liver oil	1 drop/rat/month

TABLE 2

DIETARY FACTORS THAT ALTER THE CARCINOGENICITY OF
p–DIMETHYLAMINOAZOBENZENE

Diet	Tumor incidence at	
	4½ months (end of dye feeding)	6½ months
Control (5% corn oil)	per cent 0-30	per cent 50 80
High riboflavin (5X control)	0	0
Hydrogenated coconut oil (HCNO)	0	0 8
Fatty acids of HCNO	0	0
Lauric acid (2½%)	0	0
Low fat	0	20
Egg white	0	0-14
Detergent (0.25%)	0	0
Rice bran extract (2% "Vitab") with crude casein	60--80	100 (5½ months)
Corn oil (20%)	80-100	100 (5½ months)
Low riboflavin (½X control)	60	90

TABLE 2 is a summary of some of the effects which changes in diet have on the carcinogenicity of p-dimethylaminoazobenzene. Complete protection against tumor formation was obtained if the riboflavin content of the control diet was raised to 10 mg. per kg., 5 times the level in the control diet. This is a duplication of the riboflavin-casein effect observed with the rice-carrot diet,[5] since the diet contained adequate protein. A similar protective effect has been obtained by replacing the corn oil in the diet by hydrogenated coconut oil, the fatty acids of this fat, or its chief constituent, lauric acid.[6, 7, 8] If the corn oil was replaced by glucose, the resulting low fat diet allowed only a low tumor incidence[6, 9]. The latter diets are characterized by their low content of the essential fatty acids. The typical syndrome produced by a lack in these nutrients appeared after the diet was fed for 2 to 3 months, but the prevention of this deficiency by ethyl linolate supplementation did not affect the tumor incidence[8]. The tumor incidence was also very low when the casein in the diet was replaced by egg white[10] at equivalent riboflavin levels, even though the resulting biotin deficiency syndrome was prevented by the injection of sufficient biotin. More recently, we have found that the addition of 0.25 per cent of either of two commercial synthetic detergents, known as Penetrants 4 and 7,* completely prevented the appearance of tumors by 6 months.[11] Conversely, the tumor incidence obtained with the control diet can be increased to 60-100 per cent at 4 to 4½ months and to 90-100 per cent at 5½, 6, or 6½ months, by (1) replacing the crystalline B complex in the diet by a rice bran extract characterized by low riboflavin content,[3] (2) lowering the riboflavin content of the diet to 1 mg. per kg., or (3) raising the level of corn oil in the diet to 20 per cent.[9]

TABLE 3 presents, in detail, the effect of two of these strong dietary inhibitors, riboflavin and hydrogenated coconut oil, on a carcinogen stronger than p-dimethylaminoazobenzene and a carcinogen weaker than this dye.[12, 13] m'-Methyl-p-dimethylaminoazobenzene is about twice as active as the unsubstituted dye, since the tumor incidence when 0.032 per cent of the m'-methyl compound was fed for 4 months was slightly higher than that obtained with twice the molar level or 0.060 per cent of p-dimethylaminoazobenzene. However, the inhibition obtained with riboflavin and hydrogenated coconut oil was not nearly as marked, even when the more active dye was fed at low levels, as when p-dimethylaminoazobenzene was given. Similarly, these inhibitors showed only small effects in a single experiment in which a high level (1½ molar) of a weak carcinogen, o'-methyl-p-dimethylaminoazobenzene, was employed. Hence, some of the dietary effects known for p-dimethylaminoazobenzene may be relatively specific for this dye.

A common basis has been sought for the effects produced by the several changes in the control diet on the carcinogenicity of p-dimethylaminoazobenzene. Since complete protection against tumor formation by 6 months can be effected simply by lowering the concentration of the azo dye in the diet from 0.06 to 0.03 per cent,[8] we studied both the stability of the dye in these diets and its metabolism under different dietary conditions.

---

* Branched chain alkyl sulfates manufactured by Carbide and Carbon Corporation.

TABLE 3

EFFECT OF DIET ON THE ACTIVITY OF A STRONG AND A WEAK CARCINOGEN

Percentage of dye in diet	Diet	Time dye was fed months	Liver tumors: At end of dye feeding	Liver tumors: Two months later	Liver tumors: Per cent	Negative survivors 2 months later
m'-Methyl-p-dimethylaminoazobenzene						
.048	Control	3	6/13	11/13	83	2
"	HCNO	"	6/14	12/14	86	2
"	High riboflavin	"	3/15	9/15	60	4
.032	Control	4	9/15	12/15	80	3
"	HCNO	"	5/15	8/15	53	5
"	High riboflavin	"	5/15	7/15	47	6
.026	Control	7	3/14	—	21	—
"	HCNO	"	4/11	—	36	—
"	High riboflavin	"	1/14	—	7	—
o'-Methyl-p-dimethylaminoazobenzene						
.096	Control	4	1/14	9/14	64	2
"	HCNO	"	4/13	6/13	46	4
"	High riboflavin	"	3/13	5/13	38	7

Studies showed that the control diet could be exposed at room temperature for at least 1 to 2 months without any destruction of p-dimethylaminoazobenzene or development of obvious rancidity.[7] However, once the diets containing corn oil became rancid, destruction of the dye was found. As one would expect, neither rancidity nor destruction of dye was obtained with the low fat or the hydrogenated coconut oil diets, although they are protective in nature. Kensler[11] has observed a rapid demethylation of p-dimethylaminoazobenzene in one of his freshly mixed diets. However, we have found no evidence of demethylation* prior to ingestion with any of our diets, unless they were rancid. When a corn oil diet was exposed, at room temperature, for 3 months, at least 50 per cent destruction of the dimethyl compound occurred, as well as variable amounts of demethylation to p-monomethylaminoazobenzene. No p-aminoazobenzene was detected. Perhaps it should be recalled here that the dimethyl and monomethyl compounds are, essentially, equally strong carcinogens,[16] whereas p-aminoazobenzene is non-carcinogenic.[12] Since we mix our diets for only 10-day feeding periods and even refrigerate the high fat diets, it does not seem possible to ascribe any of the dietary effects which we have observed to loss or alteration of the dye prior to ingestion. We have recently found that the demethylation occurring in the rancid diet can be duplicated, within a few days, with a more purified system, that is, p-dimethylaminoazobenzene can be demethylated to p-monomethylaminoazobenzene in the presence of autoxidizing linoleic acid with yields of up to 85 per cent and with a loss of dye amounting to approximately 5 per cent after 5 days. Small amounts of p-aminoazobenzene appeared early in the oxidation but disappeared later.

---

* See reference [15] for method.

We have also determined the levels of the 3 free azo dyes found in the body after p-dimethylaminoazobenzene was fed in these diets.[11] No significant changes were found in the amounts of these dyes in the liver or in the level of p-aminoazobenzene in the blood. The level of the latter dye in the blood would seem to be a useful indicator of absorption, since its level in the blood varies directly with the concentration of p-dimethylaminoazobenzene in the diet.[11] Likewise, the urinary levels of the amino metabolites of the dye, which are predominantly p-aminophenol and p-phenylene diamine in conjugated form, and which together account for nearly 50 per cent of the ingested dye, were unaffected by dietary changes.[17, 18] The metabolic fate of the remaining dye, and possible variation under any of these dietary conditions, are still to be determined.

Recently, we have made analyses of the liver for 3 B vitamins which have suggested a factor that appears to be common to many of the dietary effects observed. Groups of rats were fed p-dimethylaminoazobenzene in various diets for 19 weeks, and the dye-free diets for 8 weeks thereafter. Other animals received the dye-free diets for either 6 or 19 weeks, the times at which the livers of rats from each of these groups were also analyzed. The rats were kept in groups of 6 to 8 and fed weighed amounts of diet each day, so that the food intakes were kept at an average of 9 g. per rat per day. The levels of pyridoxine and biotin in the livers were determined by the growth of *Saccharomyces carlsbergensis* and *S. cerevisiae*, respectively, according to the general method of Atkin *et al.*[19] The riboflavin content was determined by the fluorometric method of Conner and Straub.[20] The choice of the latter vitamin for analysis is obvious, because of its ability to inhibit the carcinogenic activity of the dye. Similarly, pro-carcinogenic effects have been claimed for both pyridoxine and biotin. Though it has been found that a deficiency in pyridoxine retards tumor development due to p-dimethylaminoazobenzene, it was not included in TABLE 2, since the effect has not been entirely consistent and a relatively severe deficiency is needed to obtain protection.[3, 21] In contrast to a pyridoxine deficiency, the inhibitors listed in TABLE 2 allow the rat to gain weight and even consume more dye than those fed the more carcinogenic diets. The pro-carcinogenic action of biotin has been obtained by du Vigneaud *et al.*[22] and confirmed by Harris, Krahl, and Clowes.[23] In each instance, the effect was obtained with diets high in riboflavin. Thus, it seemed important to determine how the levels of these two vitamins, which have been implicated in stimulating tumor development, varied with the protective character of our diets. In FIGURE 1, the levels of pyridoxine and biotin in the livers are shown, at 19 weeks. The results are expressed as micrograms of the vitamin per gram of fresh weight. Similar results were obtained when the levels were calculated on a fat-free dry weight or on a nitrogen basis. The diets are grouped according to tumor incidence, and each bar represents the average content of at least 3 livers. Since similar data at 6 weeks were quite comparable to these, it appears that the levels were stable by that time. Administration of the dye lowered the level of each vitamin in the liver in nearly every case; the drop observed with the low fat diet was not significant. However, the amount of the decrease caused by the dye did not seem to be correlated with tumor incidence, and it should be noted that there is no evident correlation between the pyridoxine

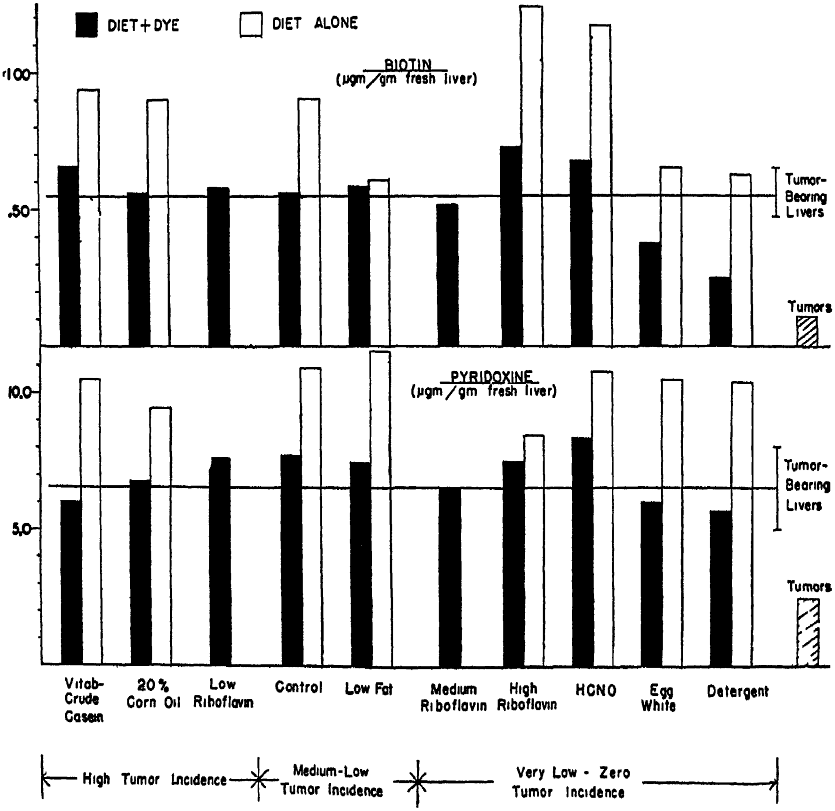
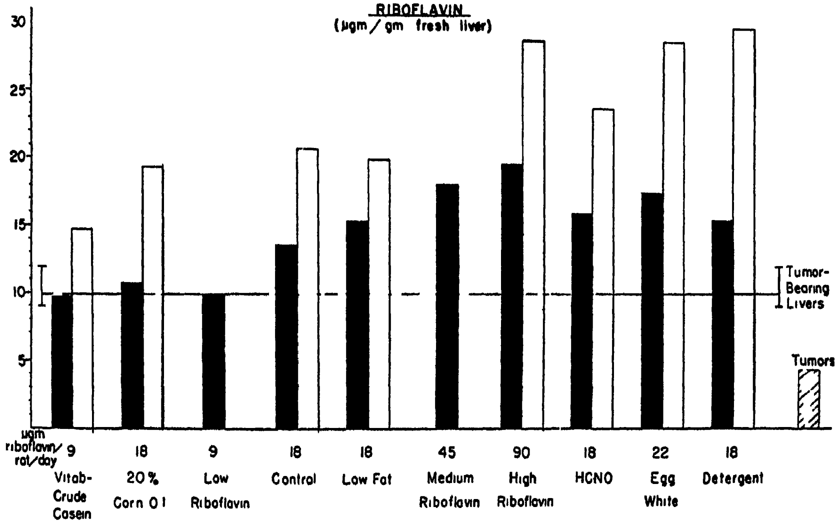


FIGURE 1. The levels of biotin and pyridoxine in the livers of rats fed p-dimethylaminoazobenzene in various diets.

or biotin contents of these livers and the protective character of the diet. The relatively low biotin content of the egg-white group was not surprising, although these animals received 2 μg. of biotin per rat 3 times each week by subcutaneous injection. This prevented any outward deficiency symptoms but did not alter the high protection offered by the diet.[10] The low level of biotin in the livers of the animals fed the control diet plus 0.25 per cent detergent is noteworthy, although these rats showed no symptoms of any deficiency, grew well, and had macroscopically normal livers at 6½ months. All of the liver tumors were lower in their contents of these two vitamins than the normal livers or the contiguous liver tissue of the host. This is in general agreement with the findings of other investigators for these and other vitamins in the type of tumor under discussion, as well as in others. In particular, the biotin levels reported here agree well with previously published values.[24] However, the pyridoxine content of the tumors given in FIGURE 1 is decidedly higher than those reported by the Texas group.[25] This may be due to differences in the pyridoxine content of the diets and the procedure used for its liberation.

In contrast to the results just discussed, the level of riboflavin in the liver correlated fairly well with the ability of a diet to protect against tumor formation. The data on this vitamin are given in FIGURE 2. Here,



Tumor Incidence, %

4½ mos	60-80	80-100	60	0-30	0	0	0	0	0	0
6½ mos	(100	100) 5½ mos	90	50-80	20	0	0	0-8	0-14	0

FIGURE 2. The levels of riboflavin in the livers of rats fed p-dimethylaminoazobenzene in various diets.

again, the dye produced a drop in the level of the vitamin in the liver that seems to be irrespective of the diet. Relatively low riboflavin levels of 9.6 to 10.6 micrograms per gram were found in the diets which permitted a high tumor incidence. The diets which gave medium to low tumor incidences produced medium riboflavin levels in the liver, of 13.4 to 15.2 micrograms per gram, respectively. The diets that offered nearly complete to complete protection at 6 months produced the highest levels of riboflavin in the liver, varying from 15.3 to 19.4 micrograms per gram. Thus, a high riboflavin level in the liver was maintained on the protective diets, even though the intake of riboflavin in several cases was relatively low. The range of the values for the livers bearing tumors was 9.0 to 11.9 micrograms per gram, with an average of 9.8 micrograms per gram. Again, the tumors were low in riboflavin, with an average of 4.4 micrograms per gram. The values given for the livers and the tumors in the high tumor incidence groups correspond to those previously published by others[5] for the rice-carrot diet, the diet that has been widely used for producing a high incidence of tumors with this dye. While no explanation is at hand for the effects produced by these diets on the level of riboflavin in the liver, it seems likely that at least part of the effect may be ascribed to intestinal synthesis of riboflavin.

Other experiments from this university support the above data on the importance of the riboflavin content of the liver to carcinogenesis by certain of the azo dyes. Thus, Griffin and Baumann[26] have found a rough correlation between the carcinogenicity of several azo dyes and their respective abilities to alter the level of riboflavin in the liver of the rat. m'Methyl-p-dimethylaminoazobenzene, the most active carcinogenic azo dye known, produced a much greater reduction of riboflavin in the liver than the very weak carcinogens, p'-methyl-p-dimethylaminoazobenzene and o-aminoazotoluene. Carcinogens of intermediate activity, such as p-dimethylaminoazobenzene and o'-methyl-p-dimethylaminoazobenzene, produced intermediate levels of riboflavin, while the non-carcinogenic azo dyes, such as p-aminoazobenzene and azobenzene, produced only a slight drop in the level of this vitamin.

In summary, it appears that the level of riboflavin in the liver of the rat is an important factor in determining the probability that a given liver will develop a tumor when p-dimethylaminoazobenzene is fed in various diets. Since other workers have not found a similar effect with spontaneous mammary tumors[27] or with tumors due to methylcholanthrene,[28] it is possible that the action of riboflavin in counteracting carcinogenesis with this dye occurs prior to the carcinogenic process, *e.g.*, through detoxication. However, enzymatic systems containing riboflavin may be involved directly in the carcinogenic process initiated by the dye. In any case, the effect of riboflavin should be a good tool in further attacks on the nature of this carcinogenic process.

BIBLIOGRAPHY

1. **Rusch, H. P., C. A. Baumann, J. A. Miller, & B. E. Kline**
   1945. Experimental liver tumors. A.A.A.S. Research Conference on Cancer, Washington, D. C.: 267-287.
2. **Miller, J. A., D. L. Miner, H. P. Rusch, & C. A. Baumann**
   1941. Diet and hepatic tumor formation. Cancer Research **1**: 699.
3. **Miner, D. L., J. A. Miller, C. A. Baumann, & H. P. Rusch**
   1943. The effect of pyridoxine and other B vitamins on the production of liver cancer with p-dimethylaminoazobenzene. Cancer Research **3**: 296.
4. **Nakahara, W., K. Mori, & T. Fugiwara**
   1939. Inhibition of experimental production of liver cancer by liver feeding. A study in nutrition. Gann **33**: 406.
5. **Kensler, C. J., K. Sugiura, N. F. Young, C. R. Halter, & C. P. Rhoads**
   1941. Partial protection of rats by riboflavin with casein against liver cancer caused by dimethylaminoazobenzene. Science **93**: 308.
6. **Kline, B. E., J. A. Miller, H. P. Rusch, & C. A. Baumann**
   1946. The carcinogenicity of p-dimethylaminoazobenzene in diets containing the fatty acids of hydrogenated coconut oil or of corn oil. Cancer Research **6**: 1.
7. **Miller, J. A., B. E. Kline, H. P. Rusch, & C. A. Baumann**
   1944. The carcinogenicity of p-dimethylaminoazobenzene in diets containing hydrogenated coconut oil. Cancer Research **4**: 153.
8. **Miller, J. A., B. E. Kline, H. P. Rusch, & C. A. Baumann**
   1944. The effect of certain lipids on the carcinogenicity of p-dimethylaminoazobenzene. Cancer Research **4**: 756.
9. **Kline, B. E., J. A. Miller, H. P. Rusch, & C. A. Baumann**
   1946. Certain effects of dietary fats on the production of liver tumors in rats fed p-dimethylaminoazobenzene. Cancer Research **6**: 5.
10. **Kline, B. E., J. A. Miller, & H. P. Rusch**
   1945. Certain effects of egg white and biotin on the carcinogenicity of p-dimethylaminoazobenzene in rats fed a sub-protective level of riboflavin. Cancer Research **5**: 641.
11. **Miller, J. A., B. E. Kline, & H. P. Rusch**
   1946. The inhibition of the carcinogenicity of p-dimethylaminoazobenzene by certain detergents and the effect of diet on the levels of azo dyes in rat tissues. Cancer Research **6**: 674.
12. **Miller, J. A., & C. A. Baumann**
   1945. The carcinogenicity of certain azo dyes related to p-dimethylaminoazobenzene. Cancer Research **5**: 227.
13. **Giese, J., C. C. Clayton, E. C. Miller, & C. A. Baumann**
   1946. The effect of certain diets on hepatic tumor formation due to m'-methyl-p-dimethylaminoazobenzene and o'-methyl-p-dimethylaminoazobenzene. Cancer Research **6**: 679.
14. **Kensler, C. J.**
   1947. Effect of diet on the production of liver tumors in the rat by N,N-dimethyl-p-aminoazobenzene. Ann. N. Y. Acad. Sci. **49**(1): 29.
15. **Miller, J. A., & C. A. Baumann**
   1945. The determination of p-dimethylaminoazobenzene, p-monomethylaminoazobenzene, and p-aminoazobenzene in tissue. Cancer Research **5**: 157.
16. **Miller, E. C., & C. A. Baumann**
   1946. The carcinogenicity of p-monomethylaminoazobenzene in various diets and the activity of this dye relative to p-dimethylaminoazobenzene. Cancer Research **6**: 289.

17. **Miller, J. A., & E. C. Miller**
1946. The metabolism and carcinogenicity of p-dimethylaminoazobenzene, and its derivatives in the rat. Abstracts of A.A.A.S. Gibson Island Research Conference on Cancer, 1946. Cancer Research **7**: 39.

18. **Miller, J. A.**
Further studies on the metabolism of p-dimethylaminoazobenzene and related compounds in the rat. (In preparation.)

19. **Atkin, L. A., A. F. Schultz, W. L. Williams, & C. N. Frey**
1943. Yeast microbiological methods for determinations of vitamins; pyridoxine. Ind. Eng. Chem., Anal. Ed. **15**: 141.

20. **Conner, R. T., & G. J. Straub**
1941. Combined determination of riboflavin and thiamine in food products. Ind. Eng. Chem., Anal. Ed. **13**: 385.

21. **Miller, E. C., C. A. Baumann, & H. P. Rusch**
1945. Certain effects of dietary pyridoxine and casein on the carcinogenicity of p-dimethylaminoazobenzene. Cancer Research **5**: 713.

22. **du Vigneaud, V., J. M. Spangler, D. Burk, C. J. Kensler, K. Sugiura, & C. P. Rhoads**
1942. The procarcinogenic effect of biotin in butter yellow tumor formation. Science **95**: 174.

23. **Harris, P. N., M. E. Krahl, & G. H. A. Clowes**
1946. The effect of liver extract, egg albumin, cystine, and cysteine upon p-dimethylaminoazobenzene carcinogenesis in rats. Abstract. Cancer Research **6**: 487.

24. **Burk, D., J. M. Spangler, V. du Vigneaud, C. Kensler, K. Sugiura, & C. P. Rhoads**
1943. Biotin-avidin balance in p-dimethylaminoazobenzene tumor formation. Cancer Research **3**: 130.

25. **Pollack, M. A., A. Taylor, & R. J. Williams**
1942. B Vitamins in Human, Rat and Mouse Neoplasms. Studies on the Vitamin Content of Tissues II: 56. University of Texas. Austin, Texas.

26. **Griffin, A. C., & C. A. Baumann**
1946. The effect of certain azo dyes upon the storage of riboflavin in the liver. Arch. Biochem. **11**: 467.

27. **Morris, H. P., & W. v. B. Robertson**
1943. Growth rate and number of spontaneous mammary carcinomas and riboflavin concentration of liver, muscle, and tumor of C3H mice as influenced by dietary riboflavin. J. Nat. Cancer Inst. **3**: 479.

28. **Strong, L. C., & F. H. J. Figge**
1946. The effect of diets containing an abundance of milk, liver, riboflavin, and xanthine on methylcholanthrene carcinogenesis. Cancer Research **6**: 466.

EFFECT OF DIET ON THE PRODUCTION OF LIVER TUMORS IN THE RAT BY N,N-DIMETHYL-p-AMINOAZOBENZENE

By C. J. KENSLER

Laboratories of the Memorial Hospital, New York, N. Y.

The production of liver tumors in rats by the administration of carcinogenic azo dyes is unique in the field of chemical carcinogenesis, in that the diets used strikingly influence tumor incidence. However, the nature of this dietary effect, whether it should be attributed to altered metabolic rate or metabolic pathway of the carcinogen, changed ability of the cell to resynthesize one or more damaged components, or to other factors, is still unknown. In view of the fact that the literature on the effect of diet on the production of liver tumors by N,N-dimethyl-p-aminoazobenzene (DMB) has been reviewed[1, 2, 3] up to 1944 and that most of the new data will be covered by other papers in this publication, I should like to discuss certain areas of this problem which our group has partially explored.

The discoveries of the carcinogenicity of the azo dyes,[4] and of the dependence of their action on diet,[5, 6] were made while using a basal ration of brown rice. At the time our experiments were started, it appeared likely that the protective influence of diet might be due to one or more factors present in liver and yeast, and that the ensuing neoplasms might be the result of an "induced or conditioned deficiency"[7] of this dietary factor or factors. Consequently, our group retained the use of this basal ration.[8] As the work progressed, it became apparent that the influence of diet was by no means simple. Some factors have been found which increase the resistance to tumor production, some enhance tumor production, and some have a dual action depending on the nature of the basal ration used.

The protective effect of liver and yeast supplements to the basal brown rice ration immediately suggested the involvement of the B complex vitamins in this process. Measurement of riboflavin and a niacin-containing coenzyme, diphosphopyridine nucleotide, showed that the concentration of these two factors was decreased in the livers of rats receiving the azo carcinogen N,N-dimethyl-p-aminoazobenzene (DMB) and, furthermore, that the riboflavin level in the livers of rats on the high tumor incidence basal ration was low even in the absence of the carcinogen.[9] The riboflavin levels were normal in livers of rats on a protective (liver

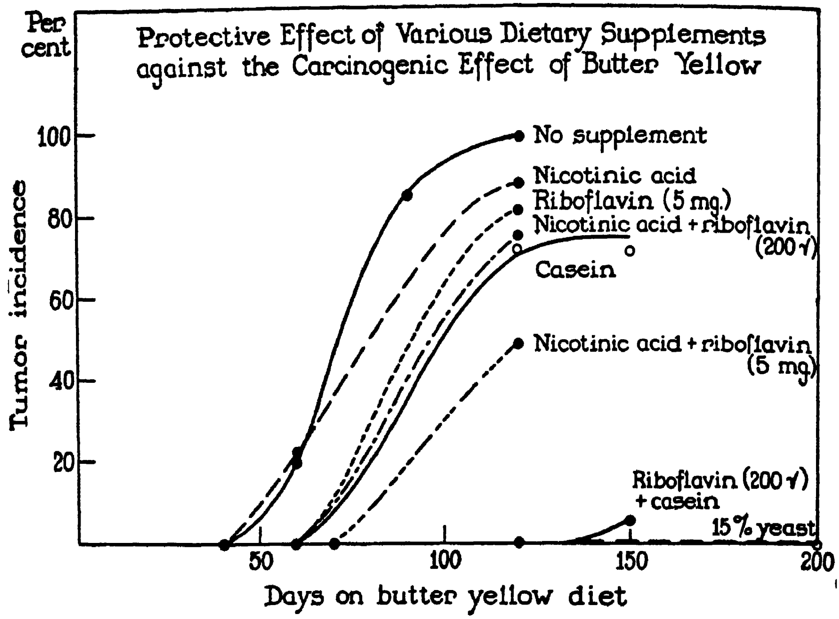


FIGURE 1. Protective effect of various dietary supplements against the carcinogenic effect of butter yellow.

or yeast) diet.[9, 10] The experiments shown in FIGURE 1 soon indicated that a riboflavin supplement, in the presence of adequate protein, increased the resistance of the rat to liver damage and tumor production by DMB.[10] The protective effect of a riboflavin supplement has been confirmed on a variety of diets and in other laboratories, and can be considered as completely established.[11, 12, 13]

The significance of this observation remains obscure. Suggestive evidence is at hand, indicating that riboflavin enzymes may be concerned in tissue protein resynthesis.[14, 15] The activity of several riboflavin-containing enzymes is decreased in the liver of riboflavin-deficient animals.[16, 17] Succinoxidase activity has also been reported decreased by one-third in livers of riboflavin-deficient rats, but not in the other tissues examined.[18] A study from our laboratory showed that the riboflavin content of rat livers is reduced when low protein diets are fed, even though adequate riboflavin is administered, and that methionine, but not cystine, tended to counteract this effect.[19] Livers from riboflavin-deficient rats also show a decreased ability to inactivate estradiol, although this is not specific for riboflavin deficiency. Thiamin deficiency and low protein intake, likewise, caused a failure of this function.[20] Some reciprocal relation between thiamin and riboflavin levels in rat livers has been indicated.[21]

Preliminary experiments were carried out in collaboration with Drs. duVigneaud and Burk, in an attempt to elucidate the protein requirement. It was found[22] that the complete replacement of the 36 per cent casein supplement by cystine, cystine and choline, or methionine did not afford much protection, whereas, when 18 per cent casein plus cystine and choline were fed, excellent protection resulted. A low riboflavin supplement also voided this protection. These experiments suggest that, although the methionine content of casein is probably important for protection, there are other needs filled by the casein (other amino acids or unknown factors) which are necessary for the maximum protective effect. The work of the Wisconsin group has indicated that protein level *per se* is probably not involved.[3]

Attempts were also made to achieve a protective brown rice diet by the use of supplements of liver and yeast fractions. Since the results with the yeast fractions are in general agreement with the liver extract experiments, only they will be presented.[23] These fractions were obtained from Dr. Clowes of Eli Lilly. The results of our tests are shown in TABLE 1. A liver extract (70 per cent alcohol) proved to be highly protective when fed at a level equivalent to 20 per cent whole dried liver (that is, 2 equivalents of a 10 per cent liver supplement, which gives excellent protection), but was practically without effect at 0.4 equivalents. A purified anti-pernicious anemia fraction gave good protection at 5 equivalents but very little at 1.3 equivalents. A riboflavin concentrate from the 70 per cent alcohol fraction, fed at 2 equivalents level, gave good protection.

TABLE 1

PROTECTION AFFORDED BY LIVER FRACTIONS

	Number equivalents 10% liver	Days on DMB diet	Number rats surviving	Number normal	Number with cirrhosis or bile duct hyperplasia	Number with hepatoma or cholangioma
Liver extract	2	139–160	20	17	3	1
	0.4	139–160	20	6	14	13
Anti P.A. fraction	5	150–216	10	6	4	1
	1.3	150–216	10	1	9	7
Flavin concentrate	2	150–270	12	8	4	1
Flavin residue	3	153–237	17	2	15	7
Flavin filtrate	3	153–237	17	15	2	1
Control diet (brown rice-carrot)	—	124–160	13	0	13	13

A fraction of this riboflavin concentrate, soluble in 70-83 per cent alcohol, at 3 equivalents, gave little protection. However, an 84 per cent alcohol-soluble fraction at 3 equivalents gave good protection. Thus, several sub-fractions of liver have been found to afford marked protection. The Wisconsin group have found that a 70 per cent alcohol extract of liver will afford protection on their low protein diet.[12] The riboflavin, biotin, and nitrogen contents of our fractions will be discussed subsequently.

Early in our work, in collaboration with Burk and duVigneaud,[24] it was noted that the addition of biotin to a brown rice protective diet decreased the protection. These experiments were confirmed with crystalline biotin on other diets. In addition, two vitamin concentrates have been found to exert a similar pro-carcinogenic effect when added to a brown rice-casein diet. Small amounts (50 mg. per 10 g. of diet) of the liver flavin filtrate soluble in 84 per cent alcohol which, as the sole supplement (500 mg. per 10 g. of diet), gave good protection, when added to an otherwise protective riboflavin casein diet destroyed the protective effect. Another supplement, 'solvamine', a vitamin concentrate from corn, when added in amounts containing 400 μg. of riboflavin › a brown rice-18 per cent casein diet failed to protect the animals as a riboflavin-casein supplement does. These data are shown in TABLE 2.

TABLE 2

EXPERIMENTS INDICATING THE PRESENCE OF A PRO-CARCINOGENIC AGENT IN TWO CONCENTRATES

Supplement	Days on diet	Number rats surviving	Number normal	Number tumor
1. 300 γ Riboflavin + 18% casein + cystine + choline	150–252	19	15	0
2. Same as (1) + 50 mg. flavin filtrate	150–210	12	3	8
3. 200 γ Riboflavin + 18% casein	150	16	10	1
4. 200 γ Riboflavin + B_1 + B_6 + P.A. + N.A. + K + 18% casein	150–230	9	8	1
5. Solvamin (containing 400 γ B_3) + 18% casein	150–200	10	0	8

On the other basal diets, the Wisconsin group have noted an increased tumor incidence when Vitab, a rice bran vitamin concentrate, was used.[3] Harris[25] has observed that liver extracts that break down protection on their basal diets can be replaced by an equivalent amount of biotin to achieve the same effect.

The analyses of our crude supplements (liver) for riboflavin, biotin (by Dr. Burk), and nitrogen, indicate that in the case of the protective fractions we are dealing with something other than a riboflavin-protein protection. These results are shown in TABLE 3. The protection afforded

TABLE 3

ANALYTICAL DATA ON VARIOUS LIVER FRACTIONS

	Amount contained in daily supplement			Protection afforded as single supplement
	Riboflavin µg.	Biotin µg.	N-total mg.	
Liver extract, 2 EQ.	130	.49	41.5	+
Liver extract, 0.4 EQ.	26	.10	8.3	—
Anti-pernicious anemia fraction, 5 EQ.	15	.07	6.0	+
Anti-pernicious anemia fraction, 1.3 EQ.	4	.02	1.5	—
Flavin concentrate, 2 EQ.	126	.30	36.6	+
Flavin filtrate, 3 EQ.	121	.35	38.6	+
Flavin residue, 3 EQ.	63	.01	33.2	—
Solvamine, 100 mg.	400	.01	6.4	

by the anti-pernicious anemia fraction which was low in both riboflavin (20 µg. per 10 g. of diet) and nitrogen (6 mg. per 10 g. of diet), is the most striking example. If the pro-carcinogenic activity of the liver flavin filtrate and solvamine is to be attributed to biotin, we shall have to assume that this effect can be caused by as little as 0.01 µg. of biotin per day.

It is quite obvious that these experiments raised more questions than they answered. Further clarification is needed in [illegible]ns of the unknown factors in these fractions which possess marked anti- and possibly pro-carcinogenic activity in DMB-feeding experiments.

TABLE 4

DEALKYLATION OF AZO DYES WHEN MIXED WITH A BROWN RICE DIET

Sample	%DMB	%MMB	%AB
1. 3% DMB in cottonseed oil—fresh	100	0	0
2. 3% DMB in cottonseed oil—4 weeks old	98.0	2.0	0
3. 0.06% DMB—brown rice—2 weeks old	69.1	27.1	3.8
4. 0.06% DMB—brown rice—1 week old on bench in beaker covered with paper	61.4	31.3	7.3
5. 0.06% DMB—brown rice—mixed and extraction started within 20 min.	84.2	13.4	2.4
6. 0.06% DMB—brown rice—heated at 90° C. for five days prior to mixing—extraction started within 20 min.	99.1	0.9	trace
7. 0.06% DEB—brown rice—1 week old	%DEB		%AB
	100		0
8. 0.06% DE-ol-B—brown rice—1 week old	%DE-ol-B		%AB
	100		0

During the course of a study of the levels of azo dyes in tissues, blood, and stomach contents of rats fed DMB, it became apparent that DMB was being demethylated to N-methyl-p-aminoazobenzene (MMB) and, to a much lesser extent, to p-aminoazobenzene (AB), when DMB in cottonseed oil was mixed with the brown rice diet. These data are presented in TABLE 4. It can be seen that the bulk of the demethylated dye recovered was the monomethyl compound MMB. It is fortunate, in terms of interpreting the nutritional experiments which have been done using the brown rice diet, that MMB is approximately equal, in carcinogenic potency,[26, 27] to DMB, and that DMB and MMB are apparently interconvertible in rat liver.[28] The fact that the riboflavin and biotin effects were first detected using the brown rice diet, and later confirmed on other diets, indicates that the demethylation occurring in the diet does not invalidate the nutritional work. It makes it desirable, however, to use other diets where the dye is stable for future nutritional work.

Recent experiments with the hydrocarbon carcinogens have indicated[29, 30, 31] that the process of carcinogenesis may be divided into three stages, *viz.*, the preneoplastic stage, a stage at which the process is reversible, and the frankly malignant stage. The concept that carcinogens can alter the cell so that other non-carcinogens can complete the conversion to the neoplastic stage, has had no known counterpart in the field of the production of liver tumors by azo dyes. Sugiura's feeding experiments[32, 33] have shown that, if DMB is fed for less than 45 days, very few animals will develop tumors, no matter whether they are then fed a protective or non-protective diet. He has also found that, if the dye is fed for 80 or more days, most of the rats will develop liver tumors, no matter what diet they are put on after DMB feeding is stopped. If a protective diet is used, cirrhosis and bile duct hyperplasia are usually not seen, although the tumors continue to grow. However, his data for an intermediate period (61-63 days) indicate that the type of diet on which

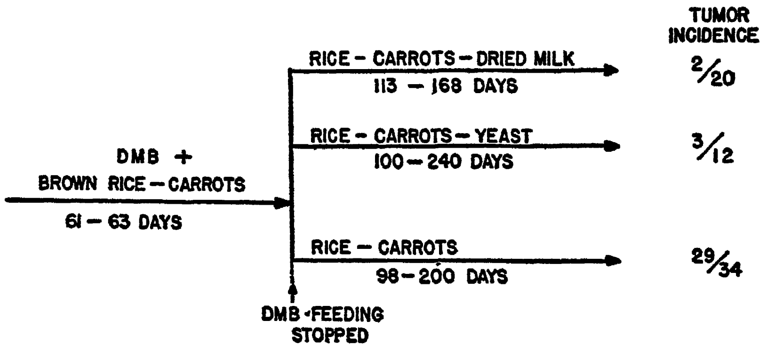


FIGURE 2. Protective influence of diet on liver tumor formation after DMB administration is stopped.

the animals are placed exerts a marked influence on the final tumor incidence. These results are summarized in FIGURE 2. It can be seen that most of the rats kept on the rice carrot diet, 29 out of 34, went on to develop tumors. However, only 2 out of 20 whose diet was supplemented with dried milk, and 3 out of 12 supplemented with yeast, developed liver tumor. In other words, we have here a suggestion that there is a stage, in the development of liver tumors by DMB, at which the outcome can be decisively influenced by diet alone if the administration of the dye is stopped. Experiments have shown that not more than a few micrograms of the methylated azo compound are present in the liver or other body tissues, except in the stomach contents when fed, at any time. This is in marked contrast to the 4000-6000 μg. that are destroyed by the rat per day. Hence, it is probable that this effect does not depend on increased destruction or detoxification of the azo dye but is dependent on the ability of the cell to restore a normal intracellular environment, provided the necessary building blocks are supplied in the diet during the critical period. Further examination of this effect should be of value in elucidating the role of diet in the final conversion of "damaged" cells into tumor cells.

Up to the present time, no chemical evidence has been obtained which would indicate that the metabolic pathway of the carcinogen DMB is altered by the use of different diets, protective *versus* non-protective, although this point has not been adequately investigated. In the rat, DMB is split at the azo linkage, and the methyl groups are removed during the course of metabolism.[34] The Wisconsin group have reported[28] that at least some of the methyl groups are removed before the azo linkage is split, and this finding has been confirmed.[35] Ethyl groups are also removed from the "non-carcinogen" N,N-diethyl-p-aminoazobenzene (DEB).[36]

We pointed out that the aromatic diamine, p-phenylenediamine (an isolated metabolite of DMB), and related diamines will, when oxidized, inhibit enzymatic activity, presumably by combining with or oxidizing SH groups of the sensitive enzyme proteins.[22, 37, 38] Potter[39] extended these observations to urease and concluded that the inactivation was due to an irreversible combination with the SH groups of this enzyme. The biological data available at the time of our publications indicated that there might be a correlation between the carcinogenic potency of the azo dye and the toxicity (enzyme inhibitory·activity) of the aromatic diamine split product derivable from it. That all azo compounds capable on paper of yielding the same split product did not possess the same carcinogenic activity was pointed out in our first paper.[37] 2-(N,N-dimethyl-p-aminophenylazo) naphthalene and methyl orange were listed. More recent evidence has shown that three closely related methyl derivatives of DMB, the 2'methyl, 3'methyl and 4'methyl compounds, all capable of yielding the same aromatic diamine, differ greatly in their carcinogenic activity. Confirmation of this finding has not appeared in the literature to date, but Dr. Sugiura's experiments with the corresponding methyl derivatives of N-methyl-p-aminoazobenzene (MMB) show the same variation in carcinogenic activity. A recent paper, by Elson and Hoch-Ligeti,[40] indicates that, in the case of urease, N,N-diethyl-p-phenylenediamine possesses approximately the same toxicity as p-phenylenediamine and less than N,N-dimethyl-p-phenylenediamine. Of the three azo compounds from which these split products would be derived which are comparable (DEB, AB, and DMB), only DMB has unequivocally been shown to produce liver tumors. It should be pointed out that no attempts to produce tumors by the administration of metabolites of carcinogens (aromatic diamines in the case of the azo compounds) have been successful.[41, 27, 26, 42]

Liver slices *in vitro* have been found to be capable of destroying DMB[35] when incubated at 37° C. (see TABLE 5). This made it appear likely that, if the carcinogen DMB were converted into toxic metabolites by liver cells, it might be possible to detect an inhibition of enzymatic activity in these slices. The only catalytic function so far investigated has been the oxidation of succinate. The succinoxidase system was selected for test, because it has been shown to be sensitive to inhibition by known and postulated metabolites of DMB, in our laboratory,[22] by Potter[43] and, recently, by Elson and Hoch-Ligeti.[40] In addition, Salter and his co-workers[44] have found that the activity of this system is decreased in the livers of rats, during the course of tumor production, by the oral administration of DMB. The activity of this system is generally quite low in tumor tissue, including liver tumors produced by DMB.

TABLE 5

DISAPPEARANCE OF AZO COMPOUNDS (50 μg.) WHEN INCUBATED WITH 200 mg. OF TISSUE SLICES AT 37.5°C. FOR 90 MINUTES

Compound	Tissue	Number of experiments	Average destroyed μg.	Range μg.
DMB	Normal liver	5	34.6	29.0–40.8
DMB	Liver tumor*	5	11.6	5.0–16.7
AB	Normal liver	5	38.4	31.3–44.8

* Classified as cholangiomas on histologic examination by Dr. S. Spitz.

As is shown in FIGURE 3, the incubation of liver slices (150 mg. wet weight) in Ringer phosphate solution, pH 7.4, containing 100 μg. of DMB which was dissolved in ethanol before addition to the medium, resulted in a decrease of from 35 to 75 per cent of the control slice succinoxidase activity at the end of a 2-hour incubation period. No effect was noted at the 20-minute point. Presumably, the delayed inhibition

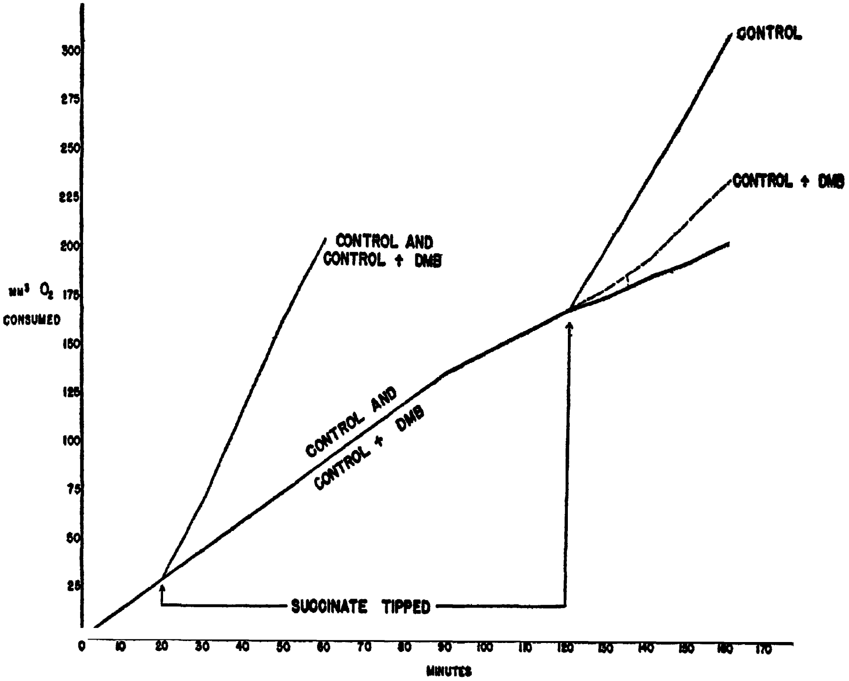


FIGURE 3. Inhibition of succinoxidase activity of rat liver slices by N,N-dimethyl-p-aminoazobenzene.

is due to the conversion of DMB into metabolites which will inhibit SH enzymes, though this point has not been proved. No inhibition of the respiration of liver slices by DMB was observed in these experiments. This is in agreement with our earlier experiments and with the findings that the *in vivo* administration of DMB has no measurable influence on the Qo_2 of liver slices[1] and that the Qo_2 of the resulting tumors is no lower than that of liver. Experiments are in progress to determine whether other closely related azo dyes, both carcinogens and non-carcinogens, will be destroyed to the same extent *in vitro* and inhibit enzymatic activity in the liver slices, and to ascertain whether these observations can be correlated with the carcinogenic potency of the azo molecules *in vivo.* It will also be of importance to ascertain whether or not the sensitivity of liver slices to this type of inhibition will be affected by the nutritional status of the rat.

The evidence available at this time neither proves nor disproves the hypothesis[22] that the azo dyes and, possibly, other carcinogens, damage cells and produce tumors because the tissue effected continuously converts the original compound into metabolites which will inactivate certain cellular enzymes, thus producing an "induced or conditioned deficiency" of this component. In the case of the production of liver tumors by the azo dyes, the preliminary evidence, shown in FIGURE 2, indicating that there is a stage in this process at which tumor incidence can be decisively influenced by diet alone, when the feeding of the carcinogen is stopped, gives further support to this hypothesis. The recent success of two groups in producing liver tumors in rats by means of diet alone,[45, 46] a chronic choline deficiency, further emphasizes the possibility that chemical carcinogens act by producing a local "induced deficiency" of the proper intensity and duration.

BIBLIOGRAPHY

1. **Burk, D., & R. Winzler**
   1944. The biochemistry of malignant tissue. Ann. Rev. Biochem. **13**: 487.
2. **Burk, D., & R. Winzler**
   1944. Vitamins and Cancer. Vitamins & Hormones II: 306. Academic Press. New York.
3. **Rusch, H. P., C. A. Baumann, J. A. Miller, & B. E. Kline**
   1945. Experimental Liver Tumors. AAAS Research Conference on Cancer 267. Science Press. Lancaster, Pa.
4. **Yoshida, T.**
   1932. Experimenteller Beitrag zur Frage der Epithelmetaplasie. Virchows Arch. Path. Anat. & Physiol. **283**: 29.
5. **Ando, T.**
   1938. Experimentelle Leber Karzinomentstehung und Getreide. I. Mitteilung. Gann **32**: 252.
6. **Nakahara, W., T. Fujiwara, & K. Mori**
   1939. Inhibiting effect of yeast feeding on the experimental production of liver cancer. Gann **33**: 57.
7. **Rhoads, C. P.**
   1942. Recent studies in the production of cancer by chemical compounds; the conditioned deficiency as a mechanism. Bull. N. Y. Acad. Med. **18**: 53.
8. **Sugiura, K., & C. P. Rhoads**
   1941. Experimental liver cancer in rats and its inhibition by rice bran extract, yeast, and yeast extract. Cancer Research **1**: 3.
9. **Kensler, C. J., K. Sugiura, & C. P. Rhoads**
   1940. Coenzyme I and riboflavin content of livers of rats fed butter yellow. Science **91**: 623.
10. **Kensler, C. J., K. Sugiura, N. F. Young, C. R. Halter, & C. P. Rhoads**
    1941. Partial protection of rats by riboflavin with casein against liver cancer caused by dimethylaminoazobenzene. Science **93**: 308.
11. **Antopol, W., & K. Unna**
    1942. The effect of riboflavin on the liver changes produced in rats by p-dimethylaminoazobenzene. Cancer Research **2**: 694.
12. **Miller, J. A., D. L. Miner, H. P. Rusch, & C. A. Baumann**
    1941. Diet and hepatic tumor formation. Cancer Research **1**: 699.
13. **Miner, D. L., J. A. Miller, C. A. Baumann, & H. P. Rusch**
    1943. The effect of pyridoxine and other B vitamins on the production of liver cancer with p-dimethylaminoazobenzene. Cancer Research **3**: 296.
14. **Sure, B., & M. Dichek**
    1941. Riboflavin as a factor in economy of food utilization. J. Nutrition **21**: 453.

15. **Kleiber, M., & T. H. Jukes**
1942. Metabolism and food utilization of riboflavin deficient chicks. Proc. Soc. Exp. Biol. & Med. **49:** 34.

16. **Axelrod, A. E., & C. A. Elvehjem**
1941. The xanthine oxidase content of rat liver in riboflavin deficiency. J. Biol. Chem. **140:** 725.

17. **Axelrod, A. E., H. A. Sober, & C. A. Elvehjem**
1940. d-Amino oxidase content of rat tissues in riboflavin deficiency. J. Biol. Chem. **134:** 749.

18. **Axelrod, A. E., V. R. Potter, & C. A. Elvehjem**
1942. Succinoxidase system in riboflavin deficient rats. J. Biol. Chem. **142:** 85.

19. **Unna, K., H. O. Singher, C. J. Kensler, H. C. Taylor, Jr., & C. P. Rhoads**
1944. Effect of dietary protein on liver riboflavin levels and inactivation of estradiol by liver. Proc. Soc. Exp. Biol. & Med. **55:** 254.

20. **Singher, H. O., C. J. Kensler, H. C. Taylor, Jr., C. P. Rhoads, & K. Unna**
1944. The effect of vitamin deficiency on estradiol inactivation by rat liver. J. Biol. Chem. **154:** 79.

21. **Singher, H. O., C. J. Kensler, H. Levy, E. Poore, C. P. Rhoads, & K. Unna**
1944. Interrelationship between thiamine and riboflavin in the liver. J. Biol. Chem. **154:** 69.

22. **Kensler, C. J., & C. P. Rhoads**
1945. Biochemical studies of chemical carcinogenesis. AAAS Research Conference on Cancer: 170. Science Press. Lancaster, Pa.

23. **Sugiura, K., C. J. Kensler, & C. P. Rhoads**
1943. Protection against the carcinogenic action of p-dimethylaminoazobenzene by liver and yeast fractions. Cancer Research **3:** 130.

24. **duVigneaud, V., J. M. Spangler, D. Burk, C. J. Kensler, K. Sugiura, & C. P. Rhoads**
1942. The procarcinogenic effect of biotin in butter yellow tumor formation. Science **95:** 174.

25. **Harris, P. N., M. E. Krahl, & G. H. A. Clowes**
1946. The effect of liver extract, egg albumin, cystine, and cysteine upon p-dimethylaminoazobenzene carcinogenesis in rats. Cancer Research **6:** 487.

26. **Miller, J. A., & C. A. Baumann**
1945. The carcinogenicity of certain azo dyes related to p-dimethylaminoazobenzene. Cancer Research **5:** 227.

27. **Sugiura, K., C. R. Halter, C. J. Kensler, & C. P. Rhoads**
1945. Observations on rats fed with compounds related to dimethylaminoazobenzene. Cancer Research **5:** 235.

28. **Miller, J. A., E. C. Miller, & C. A. Baumann**
1945. On the methylation and demethylation of certain carcinogenic dyes. Cancer Research **5:** 162.

29. **Rusch, H. P., & B. E. Kline**
1946. Further evidence for successive stages in the formation of neoplasms. Arch. Pathol. **42:** 445.

30. **Berenblum, I.**
1941. Mechanism of carcinogenesis. A study of the significance of cocarcinogenic action and related phenomena. Cancer Research **1:** 807.

31. **MacKenzie, I., & P. Rous**
1941. The experimental disclosure of latent neoplastic changes in tarred skin. J. Exp. Med. **73:** 391.

32. **Sugiura, K., & C. P. Rhoads**
1942. The effect of yeast feeding upon experimentally produced liver cancer and cirrhosis. Cancer Research **2:** 453.

33. **Sugiura, K.**
    1944. Effect of feeding dried milk on production of liver cancer by p-dimethylaminoazobenzene. Proc. Soc. Exp. Biol. & Med. **57**: 231.
34. **Stevenson, E. S., K. Dobriner, & C. P. Rhoads**
    1942. The metabolism of dimethylaminoazobenzene in the rat. Cancer Research **2**: 160.
35. **Kensler, C. J., J. W. Magill, & K. Sugiura**
    1947. Studies on the metabolism of N,N-dimethyl-p-aminoazobenzene and related compounds. Cancer Research.
36. **Kensler, C. J., J. W. Magill, K. Sugiura, & C. P. Rhoads**
    1946. Deethylation of N,N-diethyl-p-aminoazobenzene in the rat. Arch. Biochem. **11**: 376.
37. **Kensler, C. J., S. O. Dexter, & C. P. Rhoads**
    1942. Inhibition of diphosphopyridine nucleotide system by split products of dimethylaminoazobenzene. Cancer Research **2**: 1.
38. **Kensler, C. J.**
    1942. Effects of certain diamines on enzyme system; correlated with the carcinogenicity of azo dyes. Univ. Wisc. Symposium on Respiratory Enzymes: 246.
39. **Potter, V. R.**
    1942. The inhibition of sulfhydryl-containing enzymes by split products of p-dimethylaminoazobenzene. Cancer Research **2**: 688.
40. **Elson, L. A., & C. Hoch-Ligeti**
    1946. The inhibition of urease and succinoxidase by metabolic products of p-dimethylaminoazobenzene and by some related amines. Biochem. J. **40**: 380.
41. **Kinosita, R.**
    1937. Special report: Studies on carcinogenic chemical substances. Jap. Path. Soc. Trans. **27**: 665.
42. **White, J., & J. E. Edwards**
    1942. Effect of oral administration of aniline and p-aminodimethyl-aniline on the growth of the rat. Nat. Cancer Inst. J. **2**: 531.
43. **Potter, V. R., & K. P. Dubois**
    1943. Carcinogenesis and the inhibition of succinic dehydrogenase by p-dimethylaminoazobenzene derivatives. Cancer Research **3**: 133.
44. **Roskelly, R. C., N. Mayer, B. N. Horwitt, & W. T. Salter**
    1943. Studies in cancer, enzyme deficiency in human and experimental cancer. J. Clin. Invest. **22**: 743.
45. **Webster, G. T., D. H. Copeland, & W. D. Salmon**
    1942. Cirrhosis of the liver among rats receiving diets poor in proteins and rich in fat. J. Clin. Invest. **21**: 385.
46. **Engel, R. W.**
    1947. Carcinogenic effects associated with diets deficient in choline and related nutrients. Ann. N. Y. Acad. Sci. **49**(1): 49.

LEVEL OF PROTEIN INGESTION AND AN APPRAISAL IN TERMS OF PROTEIN COMPOSITION *

By Julius White, Florence R. White, and G. Burroughs Mider

National Cancer Institute, National Institute of Health, U. S. Public Health Service, Bethesda, Md., and University of Rochester, School of Medicine and Dentistry, Department of Pathology, Rochester, New York

There is no doubt that the latent period of the induction of various types of tumors in both rats and mice can be altered by a variety of dietary restrictions. Tannenbaum[1] has established that the level of caloric intake is an important factor in the formation of tumors in mice. Restriction in caloric intake resulted in a marked decrease in the incidence of every type of tumor studied. These results have been substantiated in other studies.[2, 3, 4] Other investigators[5, 6] have shown that certain vitamins of the B group can prolong the latent period of tumor formation. We have shown[7, 8, 9] that the restriction of dietary cystine markedly retards spontaneous mammary tumor formation in strain C3H mice, and induced leukemia in the strain DBA mice.

The retardation of body growth of young animals fed diets restricted either in calories or cystine does not as such explain why these animals develop fewer tumors and at a later time, than do the controls on a complete diet. Young mice maintained on a diet restricted in cystine and in which no body growth was apparent, showed no mammary tumors within a lifetime of 22 months, while the controls showed an incidence of almost 100 per cent. Yet, when the mice on the restricted diet were treated with stilbestrol, the tumor incidence rose to 44 per cent.[10] This striking difference indicated that the dietary deficiency reduced the tumor incidence by affecting certain glands of internal secretion directly and, hence, tumor genesis indirectly. Studies indicated that the mice on the restricted diets were anestrous, that the mammary tissue failed to develop in the virgin and atrophied in the breeding female, and that implantation of stilbestrol pellets in these mice resulted in growth of mammary tissue. continuous estrus, and a rise in tumor incidence (TABLE 1).

Young mice placed on an either cystine- or lysine-restricted diet may be equally retarded in growth. Painting with methylcholanthrene resulted in a low incidence of leukemia in the former group, while the incidence of leukemia in the latter was nearly as high as in the controls. These data indicate that, while dietary restriction of cystine results in a reduction of induced leukemia, a lysine-restricted diet has no apparent effect,

---

* Grateful acknowledgment is made to Dr. W. E. Heston for supplying the mice used in this investigation.

TABLE 1

EFFECT OF SUBCUTANEOUSLY IMPLANTED PELLETS OF DIETHYLSTILBESTROL UPON SPONTANEOUS MAMMARY TUMORS IN STRAIN C3H VIRGIN FEMALE MICE

Numerators = number of mice developing tumors. Denominators = number of mice dying free of tumors.
(Data taken from White & White[10])

Group of mice & diet	Effective number of mice	Mice developing tumors or dying free of tumors according to age (in months)																Tumor incidence per cent	Average age of tumors, mo.
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	22	Total		
Mice without pellet:																			
Chow	350	3/0	11/1	20/0	35/0	63/0	51/0	56/1	41/1	26/1	13/1	10/2	5/0	6/0	1/1	0/1	341/9	97.4	10.4
High cystine	42	0/0	11/1	5/0	9/0	6/0	2/0	4/0	3/0	1/0							41/1	97.6	8.3
Low cystine	45	0/0	0/0	0/0	0/0	0/0	0/0	0/3	0/2	0/0	0/1	0/0	0/2	0/1	0/0	0/0	0/9	.0	
Mice with pellet:																			
Chow	27	0/0	4/0	3/0	2/0	3/0	4/0	3/0	3/0	1/0	0/1						23/1	95.8	9.1
High cystine	28	0/1	3/1	4/0	7/0	3/0	1/0	5/0	0/0	0/0	0/0						23/2	92.0	8.4
Low cystine	40	0/0	0/4	2/3	2/2	3/1	1/1	1/1	1/1	2/1	3/1	1/1	0/1	1/4			17/1	44.7	11.1

and that the mechanism for these differences may be due to an effect on some hormonal system (TABLE 2).

TABLE 2

EFFECT OF DIET ON THE RESPONSE OF DBA MICE TO PAINTING WITH METHYLCHOLANTHRENE

(Data taken from **White, Mider, & Heston** [8])

		Mice with leukemia			Mice with sclerosis		
Diet	Number of mice	Number	Per cent	Mean latent period, days	Number	Per cent	Mean latent period, days
Dog chow	40	36	90	99	0		
High cystine	40	36	90	101	0		
Low cystine	40	4	10	127	32	80	144
High lysine	36	28	78	101	2	5.6	
Low lysine	36	26	72	104	2	5.3	

The marked difference in response by mice on these two regimes suggests that cystine played a role in the development of induced leukemia which was, perhaps, not associated with its effect as an essential amino acid for growth but with some other undetermined property.

Weissman and Schoenheimer[11] have shown that lysine labeled with N^{15} and administered to rats, can be steadily deaminated and yield its nitrogen for the formation of other amino acids. This is also true of other essential amino acids studied. However, only in the case of lysine does the reverse of this process not take place and lysine is never regenerated but must be supplied as such in the diet. Because of this unusual property of lysine, it was felt that the study of the effect of some other essential amino acid on the incidence of induced leukemia was desirable.

Accordingly, three diets were prepared in such a manner that the amount of cystine, lysine, and tryptophane, respectively, would be considerably restricted not to allow growth (increase in weight) but to warrant indefinite maintenance. These same diets, each supplemented with the amino acid in which it was restricted, permitted good growth. Another group of mice was fed a Purina dog chow diet.

TABLE 3 contains the various ingredients used in each diet. Inasmuch as the cystine-restricted diet was somewhat different from the one employed previously,[7] an additional comparison was made between the "new" and the "old" cystine-restricted diets. In our previous studies, we had used Ryzamin B #1 as part of our source of the vitamin B supplements, and liver extract (Lilly 343) to supply the remainder. However, the manufacture of Ryzamin B #1 was discontinued and the product replaced by a Ryzamin B #2 which was a preparation nearly complete in the B vitamins needed. The use of liver extract was, therefore, discontinued.

TABLE 3

Composition of the Diets*

Diet constituent	Group									
	I	IA	II	IIA	III	IIIA	IV	IVA	V	VI
	Per cent	Per cent	Per cent	Per cent	Per cent	Per cent	Per cent	Per cent	Per cent	Per cent
Labco casein vit-free	3.9	3.9	5.0	5.0	3.0	3.0				
Peroxide treated casein					7.0	7.0				
Gliaden							18.0	18.0	14.0	
Starch (corn)	55.1	54.6	66.0	65.6	60.72	60.62	53.0	52.38	57.0	
Crisco	22.0	22.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	
Cod liver oil	3.0	3.0								
Corn oil			5.0	5.0	5.0	5.0	5.0	5.0	5.0	
Inorganic salts	6.0	6.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	
Ryzamin B #1	5.0	5.0								
Ryzamin B #2			5.0	5.0	5.0	5.0	5.0	5.0	5.0	
Liver extract (Lilly)	5.0	5.0					—			
l(—)cystine		0.5		0.5						
l(+)lysine								0.62		
l(—)tryptophane						0.1				
dl-methionine					0.28	0.28				
Purina chow										100

* In addition, all diets except I, IA, and VI contained the following per kilogram of food: Natola, 150 mg.; choline chloride, 2.0 gm.; α-tocopherol, 150 mg.; and vitamin K (2-methyl 1;4-naphthoquinone), 750 mg.

Diet I : Basal cystine.
Diet IA : Basal cystine + cystine.
Diet II : Basal cystine (new).
Diet IIA : Basal cystine + cystine.
Diet III : Basal tryptophane.
Diet IIIA: Basal tryptophane + tryptophane.
Diet IV : Basal lysine.
Diet IVA: Basal lysine + lysine.
Diet V : Basal lysine (new).
Diet VI : Chow.

Weanling strain DBA mice of both sexes were used. They were allowed to eat Purina dog chow until they reached a body weight of 16 g. Then, they were distributed amongst the nine diets, so that in no case was there more than one mouse of the same sex and litter in the same group. On alternate days, the mice were painted with a 0.2 per cent ether solution of methylcholanthrene, until 20 paintings had been applied. The painting site was changed with each application, as in previous experiments.

The tryptophane-free protein was made available by treating casein with hydrogen peroxide according to the method of Toennies.[12] This process not only destroys the tryptophane but also the methionine and some of the cystine. By using 3 per cent untreated casein and 7 per cent hydrogen peroxide treated casein, a protein preparation was obtained which allowed no growth but excellent maintenance. *dl*-Methionine was added to this diet, to allow for that lost during the peroxide treatment of the casein.

The restricted lysine diet was made by using 18 per cent gliadin as a source of protein. During the course of the experiment, it became apparent that this level of gliadin permitted the mice to grow (gain weight) slightly, rather than just maintain their weight as in the cystine and tryptophane groups. Therefore, it was necessary to add, for comparative purposes, another group in which the restriction in growth on the lysine diet would approach that of the cystine and tryptophane diets. This was readily accomplished by reducing the gliadin content* from 18 to 14 per cent. With the exception of the animals on the "old" cystine-restricted and cystine-supplemented diets (Groups I and IA, TABLE 3) as well as the group on chow (Group VI), all mice received a vitamin supplement, as shown in TABLE 3. Food intake and weight changes were recorded semi-weekly. Each mouse was autopsied when evidence of leukemia was present grossly, or when the animal died from other causes. Sections were taken from each animal, and all gross observations were followed by microscopic examination of stained sections.

The results are summarized in TABLE 4. The leukemias observed were similar in gross and microscopic appearance to those previously described in dilute brown mice painted with methylcholanthrene.

The mice that ingested the "old" cystine-restricted diet (Group I) showed an incidence of 27.5 per cent leukemia, as compared with 17.1 per cent previously reported.[7] The "old" cystine-restricted diet supplemented with 0.5 per cent *l*-cystine (Group IA) showed a 92.1 per cent incidence of leukemia, which was the same as previously reported. The mean latent period for the occurrence of leukemia in these two groups was 131 and 106 days, respectively.

---

* With the group of mice on the 14 per cent gliadin diets, an additional group of mice on the high cystine diet was used as a control.

TABLE 4

EFFECT OF DIET ON RESPONSE OF STRAIN DBA MICE TO METHYLCHOLANTHRENE

Diet	Effective number of mice	Mice without neoplasm	Leukemia			Sclerosis of aorta		
			Number	Per cent	Mean latent period, days	Number	Per cent	Mean latent period, days
I	40	29	11	27.5	131	26	65.0	106
IA	38	3	35	92.1	106	1	2.6	68
II	40	19	22	55.0	113	21	52.5	121
IIA	38	3	35	92.1	97	0	0	—
III	39	6	33	84.6	136	3	7.9	128
IIIA	40	5	35	87.5	91	2	5.0	100
IV	39	4	35	89.9	124	0	0	—
IVA	40	4	36	90.0	110	0	0	—
V	38	5	33	86.3	105	0	0	—
VI	38	3	35	92.1	119	0	0	—

Diet I : Basal cystine.
Diet IA : Basal cystine + cystine.
Diet II : Basal cystine (new).
Diet IIA : Basal cystine (new) + cystine.
Diet III : Basal tryptophane.
Diet IIIA: Basal tryptophane + tryptophane.
Diet IV : Basal lysine.
Diet IVA: Basal lysine + lysine.
Diet V : Basal lysine (new).
Diet VI : Chow.

Groups II and IIA show the results obtained from animals on a diet similar to that of Groups I and IA. However, these animals received Ryzamin B #2 and other vitamin supplements instead of liver extract. The incidence of leukemia in Group II was higher (55 per cent), but considerably less than in the cystine-supplemented group (92.1 per cent). The mean latent periods were 113 and 97 days, respectively. A high incidence of sclerosis of the aorta was noted in the animals in group I (this was observed previously[7]). The mean latent period for the appearance of sclerosis was 106 days. Hence, some of these animals might have developed sclerosis before leukemia would be observed. The incidence of sclerosis in Group II was 52.5 per cent, and the mean latent period 121 days. In this group, the mean latent period of sclerosis was greater than that of the mice developing leukemia, so that leukemia could have developed before the onset of sclerosis.

Groups III and IIIA, the tryptophane-restricted and supplemented animals, respectively, showed approximately the same incidence of leukemia, which was high in both groups.

Groups IV and IVA, the lysine-restricted and supplemented animals, showed a high incidence of leukemia, about the same as the control animals on the chow diet.

Group V is the extra lysine-restricted group. This group was more restricted in lysine than was Group IV (14 instead of 18 per cent gliadin). This group was added since Group IV with 18 per cent gliadin allowed a slight gain in weight and was not comparable with Groups II and III. Group V also showed a high incidence of leukemia, being in the same range as Group III. The incidence of leukemia in Group II is approximately half that found in Groups III, IV, and V. The control animals (Group VI) ingesting chow showed an incidence of 92.1 per cent leukemia, which compares favorably with previous reports.[8]

The animals ingesting the restricted cystine, lysine, and tryptophane diets showed an average weight change, during the first 60 days of the experiment, of —0.2, +0.4, and 0.0 grams, respectively, while the animals ingesting the same diets supplemented with cystine, lysine, and tryptophane showed average respective weight changes of +4.5, +6.2, and +5.3 grams.

These results seem to indicate that the induction of leukemia in strain DBA mice by painting with methylcholanthrene can be markedly influenced by the restriction of cystine in the diet, whereas similar restrictions of the essential amino acids, lysine or typtophane, have no apparent effect. A similar reduction in leukemia has been produced by caloric restriction.[8] It is true that the mice on the restricted diets consumed less food per animal, but, based on a unit of body weight, the food consumption of the restricted and supplemented groups were equivalent. Furthermore, if one considered this as a possible effect produced by caloric restriction, it would be difficult to rationalize the marked decrease in leukemia in the cystine-restricted group and not in the lysine- and tryptophane-restricted groups.

SUMMARY

A comparative study has been made of the effect of the restriction of cystine, lysine and tryptophane, respectively, on methylcholanthrene-induced leukemia in strain DBA mice.

Each of the diets used was so restricted in one of the abovementioned amino acids that growth of young mice was prohibited but indefinite maintenance was possible. The same diets, each supplemented by the amino acid in which it was deficient, permitted good growth.

There was no significant decrease in the incidence of leukemia among the mice on diets restricted in either lysine or tryptophane.

There was a reduction in the incidence of leukemia from 92.1 per cent (control group) to 55 per cent in the group of mice whose diet was restricted in cystine.

The data indicate that, under the conditions of this experiment, cystine played a role in the development of leukemia not associated with its properties as an essential amino acid for growth, but with some other attribute not yet determined.

BIBLIOGRAPHY

1. **Tannenbaum, A.**
   1945. The dependence of tumor formation on the degree of caloric restriction. Cancer Res. **5**: 609.
2. **Visscher, M. B., Z. B. Ball, R. H. Barnes, & I. Silvertsen**
   1942. The influence of caloric restriction upon the incidence of spontaneous mammary carcinoma in mice. Surgery **11**: 48.
3. **McCay, C. M.**
   1942. Nutrition, aging, and longevity. Tr. & Studies Coll. Physicians of Philadelphia **10**: 1.
4. **White, F. R., J. White, G. B. Mider, M. G. Kelly, & W. E. Heston**
   1944. Effect of caloric restriction on mammary tumor formation in strain C3H mice and on the response of strain DBA to painting with methylcholanthrene. J. Nat. Cancer Inst. **5**: 43.
5. **Kensler, C. J., K. Sugiura, N. F. Young, C. R. Halter, & C. P. Rhoads**
   1941. Partial protection of rats by riboflavin with casein against liver cancer caused by dimethyl amino-azobenzene. Science **93**: 308.
6. **Miner, D. L., J. A. Miller, C. A. Baumann, & H. P. Rusch**
   1943. The effect of pyridoxine and other B vitamins on the production of liver cancer with p-dimethyl amino-azobenzene. Cancer Res. **3**: 296.
7. **White, J., & G. B. Mider**
   1941. The effect of dietary cystine on the reaction of dilute brown mice to methylcholanthrene. J. Nat. Cancer Res. **2**: 95.
8. **White, J., G. B. Mider, & W. E. Heston**
   1944. Effect of amino acids on the induction of leukemia in mice. J. Nat. Cancer Inst. **4**: 409.
9. **White, J., & H. B. Andervont**
   1943. Effect of a diet relatively low in cystine on the production of spontaneous mammary gland tumors in strain C3H female mice. J. Nat. Cancer Inst. **3**: 449.
10. **White, F. R., & J. White**
    1944. Effect of diethylstilbestrol on mammary tumor formation in strain C3H mice fed a low cystine diet. J. Nat. Cancer Inst. **4**: 413.
11. **Weiseman, N., & R. Schoenheimer**
    1941. The relative stability of 1 (—) lysine in rats studied with deuterium and heavy nitrogen. J. Biol. Chem. **140**: 779.
12. **Toennies, G.**
    1942. The oxidative conversion of casein in protein free of methionine and tryptophane. J. Biol. Chem. **145**: 667.

CARCINOGENIC EFFECTS ASSOCIATED WITH DIETS DEFICIENT IN CHOLINE AND RELATED NUTRIENTS *

By R. W. ENGEL, D. H. COPELAND, AND W. D. SALMON

Laboratory of Animal Nutrition, Alabama Polytechnic Institute, Auburn, Alabama

INTRODUCTION

There is ample evidence that hepatic cirrhosis may be produced in experimental animals by a deficiency of choline.[1-4] The importance of this is emphasized by the common association of hepatic cirrhosis and primary cancer of the liver, in man.

In 1946, Copeland and Salmon[5] reported the occurrence of neoplasms in 40 of 69 rats restricted to a choline-deficient diet for 8 to 16 months. No neoplasms occurred in 19 litter-mate control rats fed, for comparable periods, on the same diet, supplemented with 20 mg. of choline chloride per rat daily.

It should be noted that, in 1942, Webster[6] reported the occurrence of neoplasms in 4 of 20 rats that received an addition of cystine to a diet containing 8 per cent casein and 40 per cent fat. Since spontaneous neoplasms did not occur in animals of the same age and strain subsisting on similar diets without added cystine, Webster suggested a relation between dietary cystine and neoplasms. Although the diets used by Webster were not supplemented with choline, 2 of the 4 rats developing neoplasms received 50 mg. of betaine hydrochloride per rat daily. This should have provided an adequate labile methyl supply.

In 1945, a series of experiments was begun, in this laboratory, to determine whether neoplasms could be produced by other choline-deficient basal diets. Corn grits were used in some of the diets, in view of Salmon's observation that this ingredient accelerated growth in rats fed diets low in casein.[7] These experiments are described in the following pages.

PROCEDURE

The composition of the diets used is shown in TABLE 1. The dry ingredients were mixed thoroughly. The following accessory nutrients were mixed into each kilogram of diet: 50 mg. alpha-tocopherol, 5 mg. beta-carotene, 0.125 mg. calciferol, 2 mg. each of thiamin and pyridoxine, 4 mg. of riboflavin, 10 mg. of calcium pantothenate, 20 mg. of nicotinic acid, and 200 mg. of *i*-inositol. The diets were stored at 1° C. and

---

* Published with the approval of the Director of the Alabama Agricultural Experiment Station. This work was supported, in part, by a grant from The Nutrition Foundation, Inc., New York, N. Y.

TABLE 1

PERCENTAGE COMPOSITION OF THE DIETS

Ingredient	Diet No. 46	46E†	47	48	49	50
Degerminated corn grits	40	40			40	
Water-extracted casein	9	4.5	12	12	12	15
l-Cystine	0.3	0.3	0.3	0.3	0.3	0.3
Sucrose	31.7	36.2	82.7	68.7	28.7	65.7
Lard	15	15		15	15	15
Salts 5*	4	4	4	4	4	4
Corn oil			1.0			

* J. Nutr. 33: 155. 1947.

† One-half of the animals receiving diet 46 were changed to diet 46E at the end of 6 months, or when mature body weight had been reached.

replenished at approximately 10-day intervals. Animals that served as controls were fed the same diets, containing 2 gm. of choline chloride per kilogram.

Weanling rats, 23 days of age and weighing approximately 45 gm., were housed individually in screen-floored cages and were fed daily. Food jars were replaced weekly. The animals were from a strain of hooded rats developed, by inbreeding, from stock originally obtained from Steenbock, at the University of Wisconsin, and propagated in this laboratory since 1928.

Since it was anticipated that losses would occur from hepato-renal damage in the animals receiving the choline-deficient diets, the kidneys were palpated twice daily during the early phases of the experiment. Whenever kidney enlargement was noted, choline chloride was administered in aqueous solution by stomach tube. This therapy was repeated at 8-hour intervals until the first indications of recovery were apparent, as determined by reduction in kidney size. From 2 to 9 doses of 20 mg. each were required. Acute hepato-renal damage appeared most commonly between the first and second weeks of the experiment, but subsequent damage was not infrequent and, in some cases, recurred as late as 10 to 12 weeks after the start of the experiment. This method of administering choline during the early phases of the experiment was adopted because it proved to be a considerably more effective means of keeping losses at a minimum than did choline feeding or injection.

Autopsy material was preserved in Bouin's fixative, and sections of tissue stained with hematoxylin and eosin were, routinely, examined microscopically. The animals were sacrificed by carotid bleeding when death appeared imminent or, if death had occurred, necropsy material was obtained within a few hours *post-mortem.*

These experiments are still in progress. The present report includes data from 18 experimental animals that succumbed or were sacrificed between the fifth and eleventh months of the experiment.

RESULTS

Growth and Symptoms. Average mature body weights of the animals varied, depending upon the diet fed. For all diets used, the inclusion of choline chloride in the diet resulted in mature body weights 100 to 150 gm. higher than were obtained with the corresponding choline-deficient diets. Animals receiving diet 46 attained maximum body weights of about 400 gm., whereas the same diet without corn grits (diet 48) resulted in mature body weights of about 300 gm. Increasing the protein content of either of these diets by increasing the casein 3 per cent (diets 49 and 50) increased the final body weight only slightly. The low-fat diet (diet 47) produced slightly poorer growth than the same diet with 15 per cent lard. The animals fed diet 46E, from the sixth month to the end of the experiment, accumulated large amounts of fluid in their body cavities as well as in subcutaneous tissues.

In general, the animals receiving the choline-deficient diets grew only moderately well during the first 6 months of the experiment. The first symptom of chronic choline deficiency was a general unthrifty appearance and decline in appetite. This was followed by moderate weight loss. In some cases, there was temporary recovery, which was usually followed by moderate to rapid failure. In those animals which had lung involvement, *râles* and hemorrhagic rhinitis were consistent symptoms during the latter 6 to 8 weeks of the experiment. Labored respiration was common in animals with severe lung damage or with hydrothorax.

Pathology. The pathological conditions observed in the 18 animals that have completed the experiment are summarized in TABLE 2. Neoplasms of one or more types were present in 14 of the animals. Neoplastic changes in the lungs were the most consistent findings, occurring in 10 cases. Grossly, the involved lungs contained nodules of various sizes up to 1 cm. in diameter. Some of the nodules were filled with pus and were pliable, with purulent extrusion occurring upon incision. The smaller nodules were firm and usually densely cellular. Involvement ranged from part of a single lobe in some cases to practically the whole organ in others. Neoplastic changes most commonly seen were primary medullary in character and originated in the bronchial epithelium. Metaplasia of bronchial epithelium into stratified squamous type was frequently observed. Near and surrounding areas of metaplasia were whorls and cords of epithelial-type cells which frequently tended toward stratification. Mitotic figures were present. The inflammatory process, which frequently was present in these lungs, was characterized by infil-

TABLE 2

SUMMARY OF OBSERVED PATHOLOGICAL CHANGES IN TISSUES OF CHRONIC CHOLINE-DEFICIENT RATS

Rat No.	Diet No.	Sex	No. days on experiment	Lesions								
				Lung neoplasm	Liver neoplasm	Liver cirrhosis	Lymph involvement	Sarcoma	Hemangioma	Pancreas neoplasm	Pancreas centroacinar cell proliferation	Bladder carcinoma
9284	46	F	151	x					x	x		
9320	46	M	199	x								
9283	46	M	269					x				
9376	46	M	272									
9290	46	M	296	x		x					x	
9294	46	M	325	x								
9280	46	F	332	x		x		x	x			
9358	46E	M	214	x								
9360	46E	F	246		x	x	x		x			
9297	46E	F	247	x			x		x			
9366	46E	F	275	x	x	x				x		
9317	46E	F	294		x	x				x		
9299	46E	M	305			x	x					
9331	47	M	263									
9406	48	M	213	x		x	x	x		x		x
9349	48	F	230	x								
9395	49	F	213									
9410	50	M	238		x	x					x	

tration of blood cell elements with varying density, sloughing of alveolar cell linings, bronchiectasis and, in some cases, marked atelectasis. An increase in fibroid tissue was found in a few cases.

Nine of the 18 animals had severe "hobnail" liver cirrhosis. In 4 of these cases, liver neoplasms that were adenomatous in character were found. These growths were composed of cords or acini-like groups of cells with dense nuclei and distinct cytoplasm. These cells had a striking resemblance to liver parenchyma. In 2 cases, neoplastic cells of a reticular type, forming no distinct structural arrangement, were located primarily in interlobular areas. These cells seemed to be epithelial in origin. Numerous mitotic figures were present.

Sarcomas were found in 3 of the 18 rats. In 2 animals, these neoplasms were small in size and were located subcutaneously in the abdominal region. In the third animal, a primary sarcoma developed very rapidly in the antero-lateral aspect of the thigh, invaded the muscle tissue extensively, and metastasized along the sciatic vessels to form two secondary smaller growths. The primary growth outgrew its blood supply, so that its center was a soft mass of decaying cellular material.

Three different types of pathological changes were observed in the pancreas of 6 of the animals. An adenomatous type growth was present in 4 cases in which the organ was considerably enlarged and the neoplastic tissue replaced about one-third of the organ. Numerous mitotic figures were present, and the new growth seemed to originate from pancreatic cells, but did not form true acinar structures. Fibrous tissue proliferation was also observed in these cases. Centro-acinar cell proliferation was observed in two cases. This type of damage was less extensive than that just described. In the pancreas of one animal, a centralized growth of fibrotic tissue resembling a fibrosarcoma had replaced about one-fifth of the normal tissue.

One animal had extensive urinary bladder damage, which was the apparent cause of rapid failure and death. Papillomatous epithelial growths were prominent. In one area, this type of growth had infiltrated the muscular coats. Numerous mitotic figures were present, and typical carcinomoid tissue had replaced the transitional epithelial cells in this area.

Hemangio-endotheliomas in the mesenteric and subcutaneous fat were found in 4 of the 18 animals. Three of these were extremely cellular, contained a large number of small vessels, and exhibited numerous mitotic figures. One growth was classified as a cavernous hemangioma and consisted of large islands of blood surrounded by shelves of tissue composed of cells having spindle-shaped nuclei.

Lymph gland involvement was noted in 4 of the 18 animals. The pathology consisted of grossly apparent swelling and reddening of the cisternal, cervical, and thoracic nodes, and was characterized, microscopically, by the infiltration of abnormal cells of three types. The cells in some cases were reticular with stellate cytoplasm; in other cases, they resembled plasma cells and replaced about one-third of the normal tissue; in still others the cells resembled liver parenchyma.

There have been no abnormalities noted among the 20 control animals receiving the choline-supplemented diets, except for obvious signs of lung involvement of the inflammatory type in one case. This animal died, and autopsy revealed a severe inflammatory process involving most of the lungs.

DISCUSSION

The pathological material available at present does not justify conclusive remarks concerning the efficacy of the various diets used. However, certain observations deserve mention. Reduction of the dietary casein from 9 to 4.5 per cent (diet 46E) appeared to increase damage to the liver. In general, it appears that none of the diets used in the present study were as effective for producing extensive liver damage as was the diet employed in the previous study.[5] On the other hand, the incidence of neoplasms in other tissues appears to be higher than was observed previously.

Of interest in connection with the results obtained in this laboratory is the report of Gilbert and Gillman.[8] They observed a multiplicity of pathological changes, including extensive liver and lung damage, in rats fed diets of corn meal and fermented milk. Moreover, Gillman and Gillman[9] have suggested that there is a definite relationship between the high incidence of cirrhosis and primary liver carcinoma on the one hand, and repeated injury to the liver from acute and chronic malnutrition among adolescent and young adult humans on the other.

Although Gyorgy and associates[10] reported some definite but irregular protection with choline and cystine in butter yellow-induced liver tumors in rats, the majority of available data indicate no beneficial effect of dietary choline in this type of induced neoplasm.[11-14]

Also of interest are the reports in the older literature, by Robertson and Burnett[15] and Werner[16]. The former workers observed that injection of aqueous emulsions of lecithin diminished the tendency toward metastases, retarded metastatic growth, and retarded primary tumor growth. Werner observed that continuous injection of choline borate reduced the size of tumors in rats. Recently, Beard[17] has reported that ingestion of choline chloride in the drinking water resulted in increased disappearance of intra-abdominally implanted Emge sarcoma in rats.

SUMMARY

Choline-deficient diets, of different composition from those previously used in this laboratory, were fed to rats for prolonged periods. Neoplasms of one or more types were observed in 14 out of 18 rats fed these diets

for 5 to 11 months. No neoplasms were observed in control animals fed the same diets supplemented with 0.2 per cent choline chloride.

BIBLIOGRAPHY

1. **Gyorgy, P., & H. Goldblatt**
1941. Experimental production of dietary liver injury (necrosis, cirrhosis) in rats. Proc. Soc. Exp. Biol. & Med. **46**: 492.

2. **Lowry, J. V., F. S. Daft, W. H. Sebrell, L. L. Ashburn, & R. D. Lillie**
1941. Treatment of dietary liver cirrhosis in rats with choline and casein. Pub. Health Rep. **56**: 2216.

3. **Blumberg, H., & E. V. McCollum**
1941. The prevention by choline of liver cirrhosis in rats on high fat, low protein diets. Science **93**: 598.

4. **Engel, R. W.**
1943. Liver cirrhosis and choline. Fed. Proc. **2**: 62.

5. **Copeland, D. H., & W. D. Salmon**
1946. The occurrence of neoplasms in the liver, lungs, and other tissues of rats as a result of prolonged choline deficiency. Am. J. Path. **22**: 1059.

6. **Webster, G. T.**
1942. Cirrhosis of the liver among rats receiving diets poor in protein and rich in fat. J. Clin. Invest. **21**: 385.

7. **Salmon, W. D.**
1947. Some physiological relationships of protein, fat, choline, methionine, cystine, nicotinic acid and tryptophane. J. Nutrition **33**: 155.

8. **Gilbert, C., & J. Gillman**
1944. Diet and disease in the Bantu. Science **99**: 398.

9. **Gillman, T., & J. Gillman**
1945. Hepatic damage in infantile pellagra. J.A.M.A. **129**: 12.

10. **Gyorgy, P., E. C. Poling, & H. Goldblatt**
1941. Necrosis, cirrhosis and cancer of liver in rats fed a diet containing p-dimethylaminoazobenzene (butter yellow). Proc. Soc. Exp. Biol. & Med. **47**: 41.

11. **Jacobi, H. P., & C. A. Baumann**
1942. Choline in tumor-bearing animals and a choline-like effect of butter yellow. Cancer Res. **2**: 175.

12. **White, J., & J. E. Edwards**
1942. Effect of supplementary methionine or choline plus cystine on the incidence of p-dimethylaminoazobenzene-induced hepatic tumors in the rat. J. Nat. Cancer Inst. **3**: 43.

13. **Miner, D. L., J. A. Miller, C. A. Baumann, & H. P. Rusch**
1943. The effect of pyridoxine and other B vitamins on the production of liver cancer with p-dimethylaminoazobenzene. Cancer Res. **3**: 296.

14. **Miller, E. C., & C. A. Baumann**
1946. The carcinogenicity of p-monomethylaminoazobenzene in various diets and the activity of this dye relative to p-dimethylaminoazobenzene. Cancer Res. **6**: 289.

15. **Robertson, T. B., & T. C. Burnett**
1913. Influence of lecithin and cholesterol upon the growth of tumors. J. Exp. Med. **17**: 344. (Abstracted in Chem. Abst. **8**: 1468. 1914.)

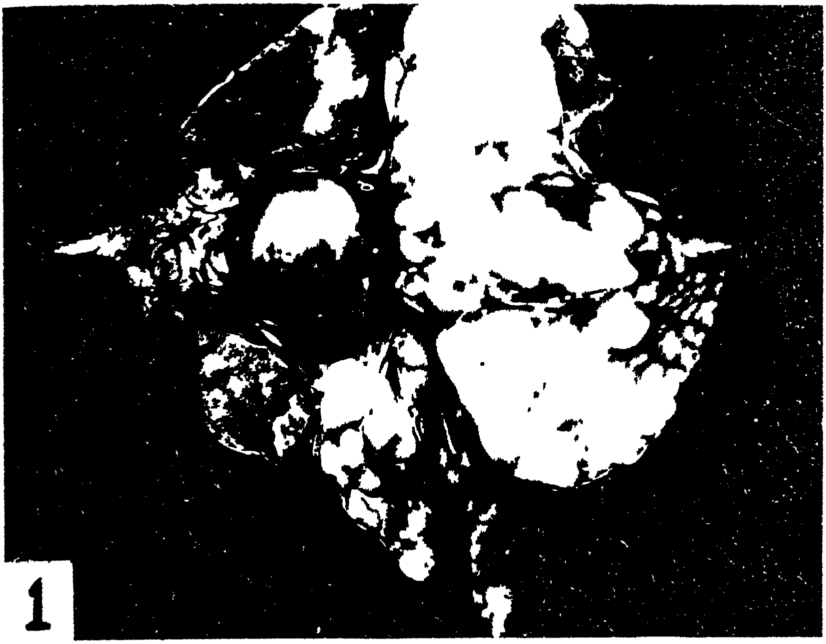
16. **Werner, R.**
1912. Chemical imitation of active rays and chemotherapy of cancer. Med. Klinik **8**: 1160. (Abstracted in Chem. Abst. **7**: 157. 1913.)

17. **Beard, H. H.**
1944. The effect of penicillin and choline upon the appearance, growth, and disappearance of the Emge sarcoma in rats. Exp. Med. Surg. **2**: 286.

PLATE 2

FIGURE 1. Lung exhibiting nodular growths of various sizes.

FIGURE 2. Metaplasia of epithelium lining a bronchiectatic cavity. The epithelium exhibits a transition from simple columnar to typical stratified squamous. This figure also illustrates the possible bronchiogenic origin of the underlying neoplasm. Hematoxylin and eosin stain. X250.



1

2

ENGEL ET AL: CHOLINE AND CANCER

ENGEL *ET AL.*: CHOLINE AND CANCER

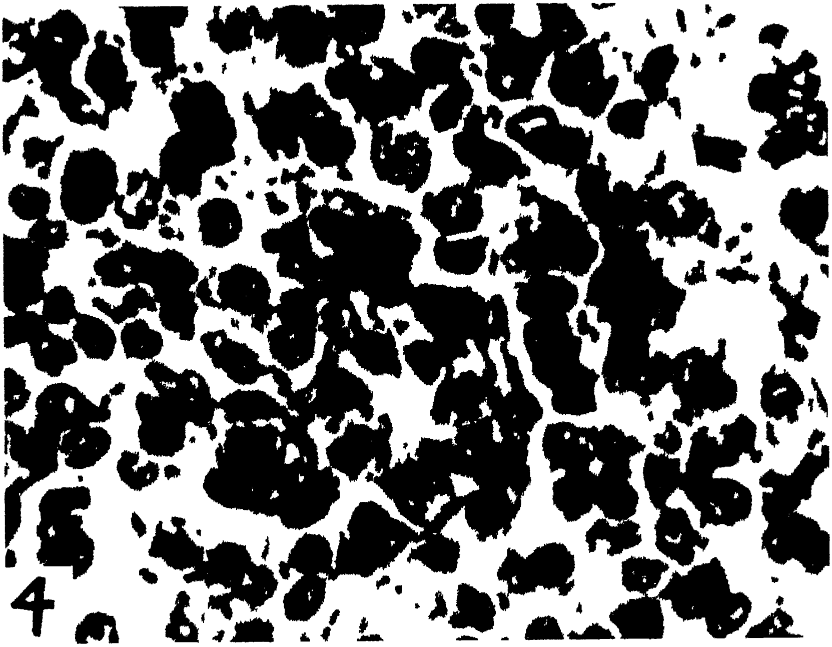
PLATE 3

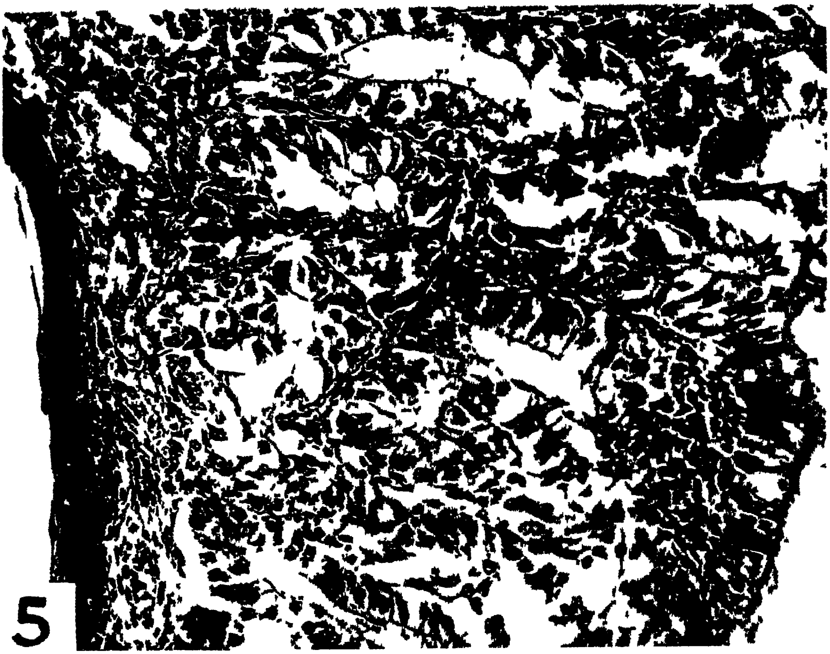
FIGURE 3. Extensive area of neoplasm of the nature of a primary carcinoma of the lung, underlying bronchial epithelium. The cells are cylindrical in type, with distinct nucleoli. They seem to fill the alveolar space. Hematoxylin and eosin stain. X510.

PLATE 4

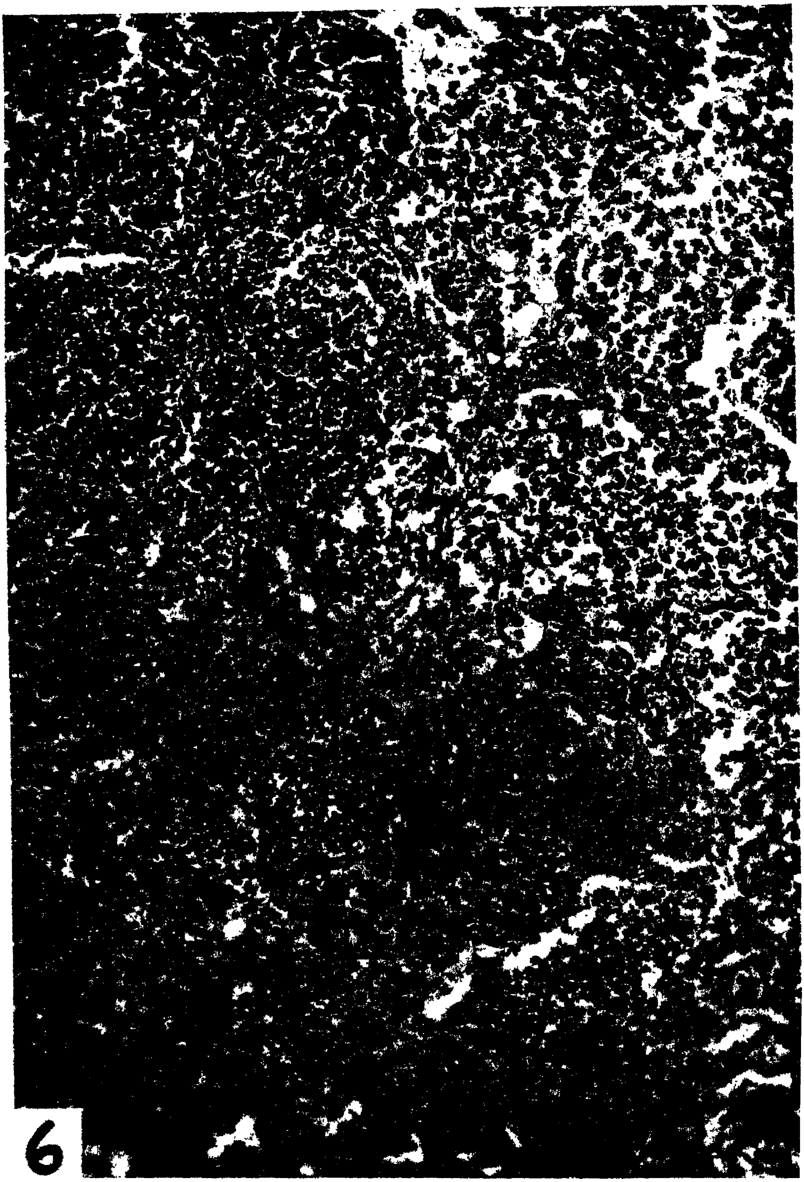
FIGURE 4. Small area of FIGURE 3 enlarged to illustrate the nature of the cells of this neoplasm. Hematoxylin and eosin stain. X1000.

FIGURE 5. Metaplasia of epithelium lining a bronchiectatic cavity, and an underlying neoplasm of an adenomatous nature. Hematoxylin and eosin stain. X510.


4


5

ENGEL *ET AL*: CHOLINE AND CANCER



6

ENGEL *ET AL.*: CHOLINE AND CANCER

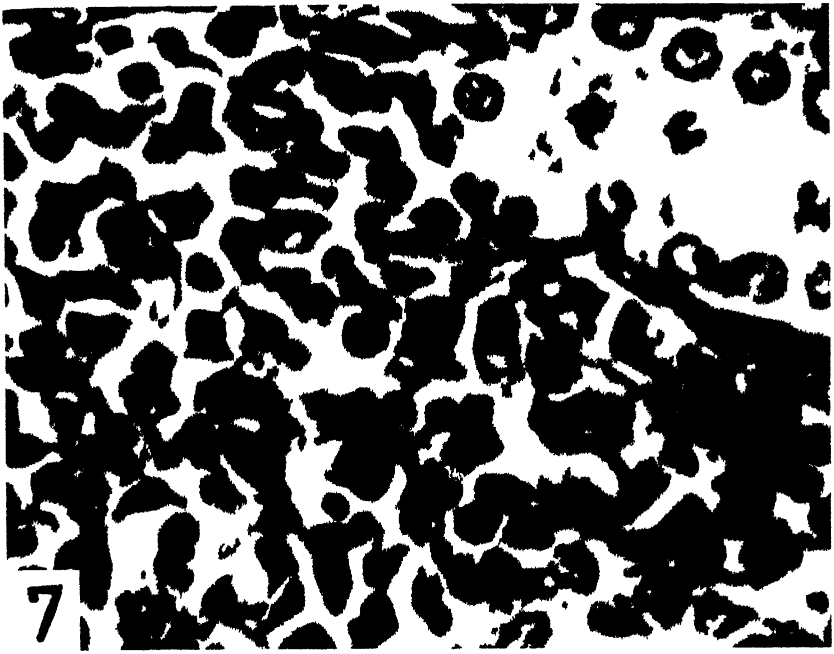
PLATE 5

FIGURE 6. Extensive area of lung showing dense inflammatory exudate and areas of granulation tissue. Hematoxylin and eosin stain. X250.

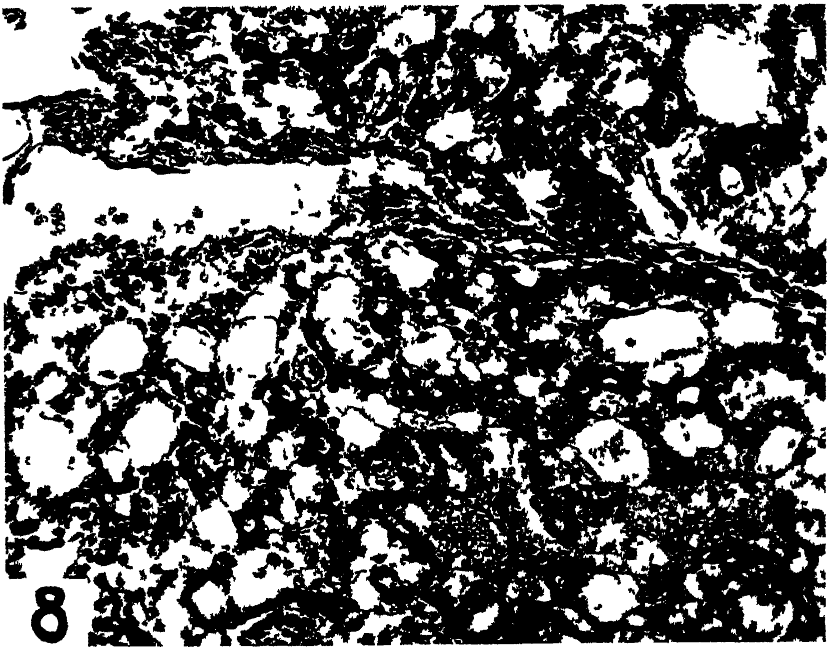
PLATE 6

FIGURE 7 Lung growth consisting of anaplastic cells with dark staining nuclei and indefinite cytoplasm The cells in this type of growth had a characteristic medullary arrangement Hematoxylin and eosin stain X1000

FIGURE 8 Neoplasm of the nature of a hemangio endothelioma, showing the vascular nature and the type and arrangement of the cells Growths of this nature were found in the mesentery and, subcutaneously, in the inguinal region Hematoxylin and eosin stain X260

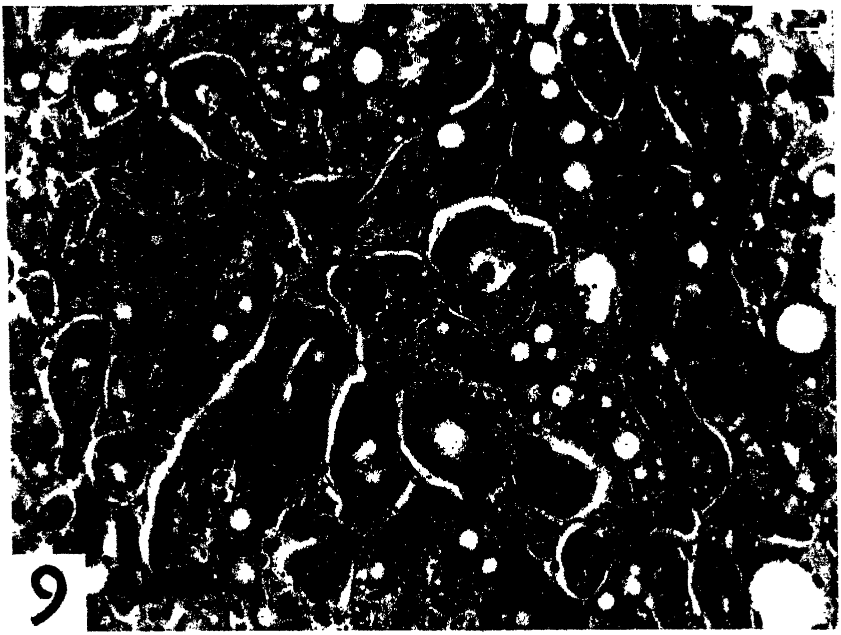


7

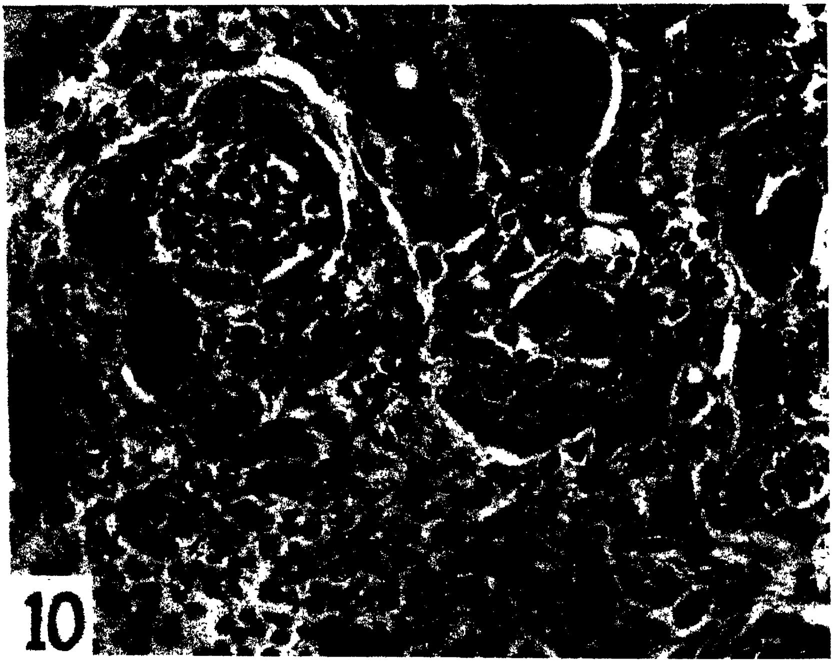


8

ENGEL *ET AL* · CHOLINE AND CANCER



9



10

ENGEL *ET AL.*: CHOLINE AND CANCER

PLATE 7

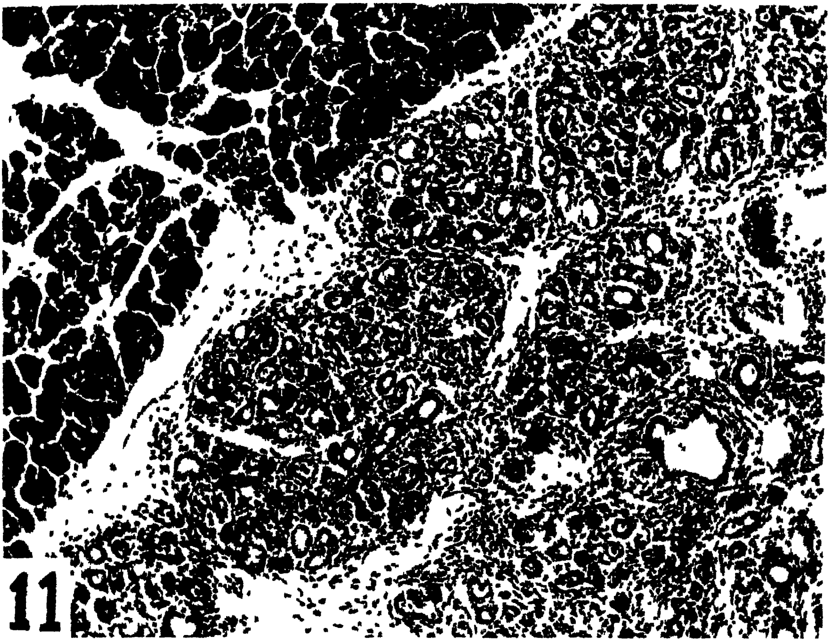
FIGURE 9. Section of the liver illustrating the possible origin of adenocarcinoma from liver parenchyma. Hematoxylin and eosin stain. X250.

FIGURE 10. Neoplasm developing in interlobular areas in the liver. Hematoxylin and eosin stain. X475.

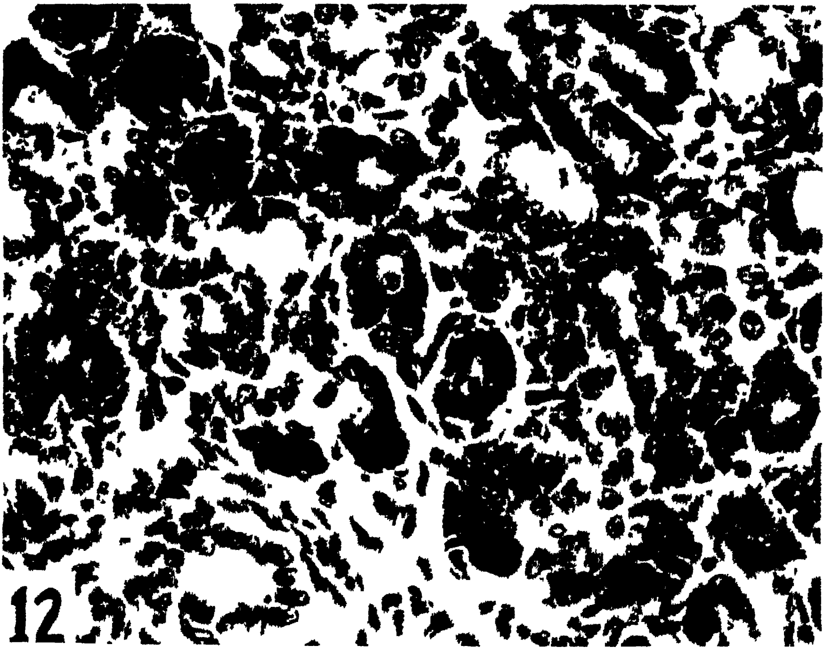
PLATE 8

FIGURE 11. Neoplasm in the pancreas. The cells of the epithelium lining the cystic spaces are apparently derived from parenchymal cells, but typical acinar structures failed to develop. Some infiltration of blood cell elements and fibroblasts is present. Hematoxylin and eosin stain. X120.

FIGURE 12. A small area of neoplasm of pancreas enlarged to show the type and arrangement of the cells. Hematoxylin and eosin stain. X475.

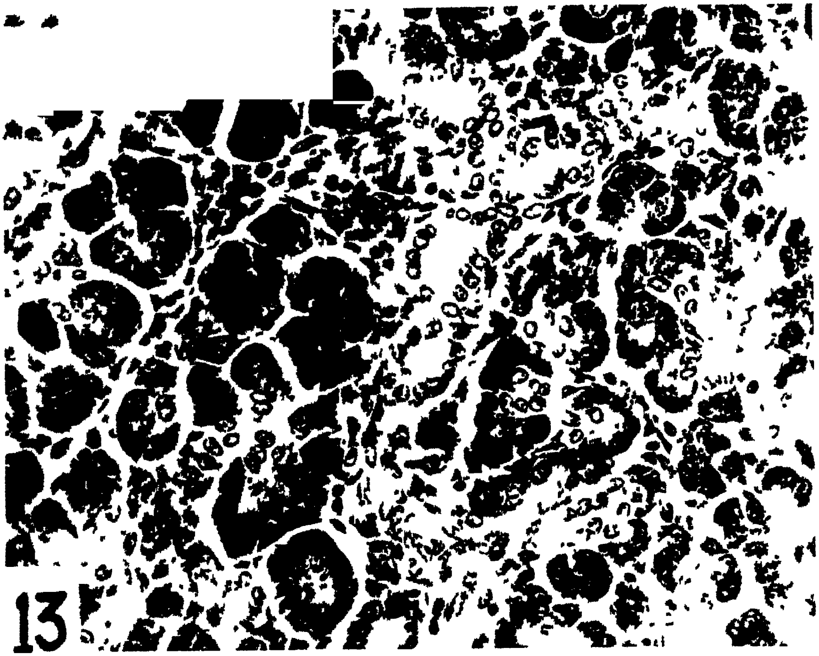


11

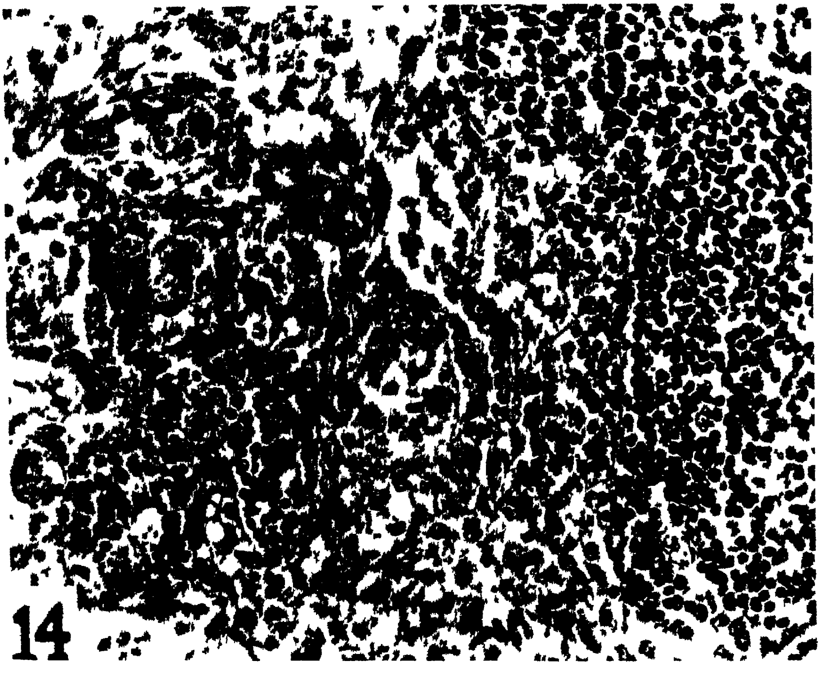


12

ENGEL *ET AL.*: CHOLINE AND CANCER



13



14

PLATE 9

FIGURE 13. Area of pancreas showing excessive proliferation of centro-acinar cells Hematoxylin and eosin stain X390.

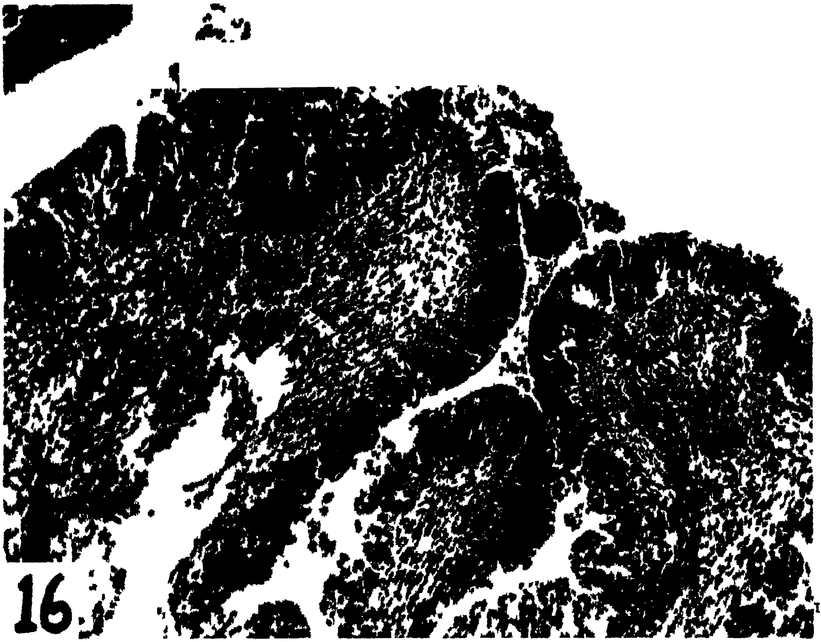
FIGURE 14 Lymph gland showing an abundance of reticular type cells of peculiar design found in medulla and cortex. Hematoxylin and eosin stain. X390.

PLATE 10

FIGURE 15 Deep pelvic dissection to illustrate urogenital tract This bladder (upper central was extremely enlarged One large ulcer was present near its base, and a growth of the nature of primary carcinoma was found at the apex

15

ENGEL ET AL.: CHOLINE AND CANCER



16

17

ENGEL ET AL.: CHOLINE AND CANCER

PLATE 11

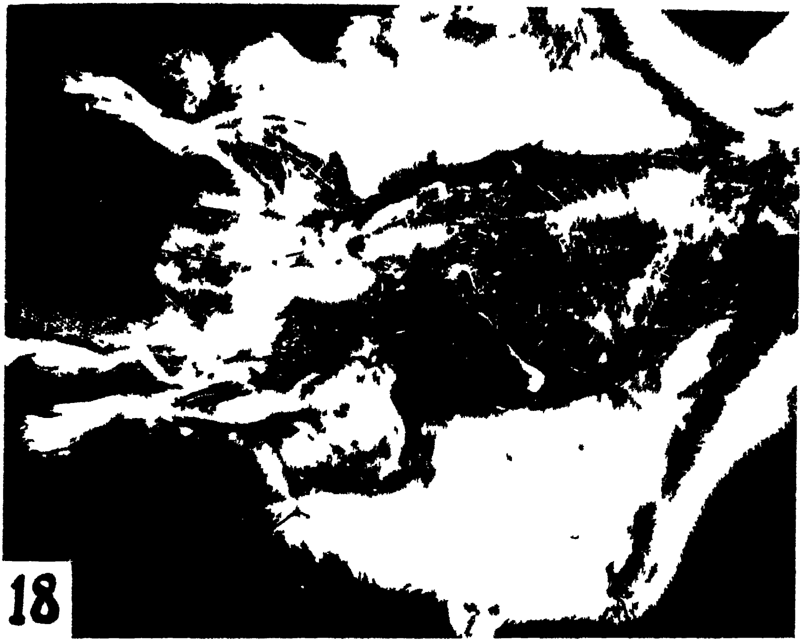
FIGURE 16. Hyperplasia and papillomata of bladder epithelium. Hematoxylin and eosin stain. X120.

FIGURE 17. An area which has the cellular appearance of a carcinoma of the bladder. Note the type and arrangement of the cells. Hematoxylin and eosin stain X330.

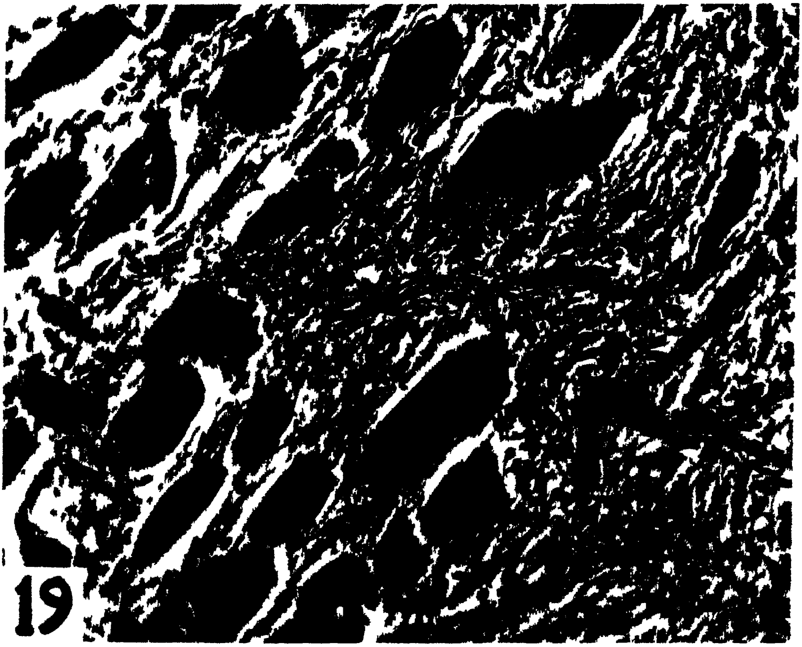
PLATE 12

FIGURE 18 Gross dissection of rat, showing a large sarcoma on the thigh

FIGURE 19 Infiltration of muscle tissue by sarcoma The cells are arranged in planes and whorled in different directions Hematoxylin and eosin stain X300

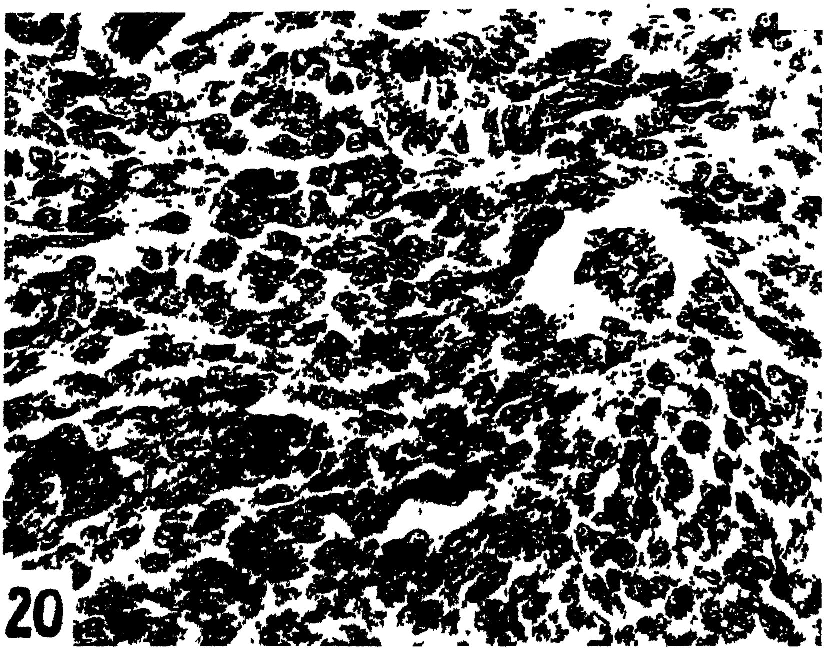


18



19

ENGEL *ET AL.*: CHOLINE AND CANCER



20

21

ENGEL *ET AL.*: CHOLINE AND CANCER

PLATE 13

FIGURE 20. A sarcoma which has extensively invaded and nearly replaced muscle. Hematoxylin and eosin stain. X500.

FIGURE 21. Enlarged area of the sarcoma, to illustrate the nature of the cells. Mitotic figures are very numerous in these growths. Hematoxylin and eosin stain. X1000.

THE MAMMARY TUMOR MILK AGENT *

By JOHN J. BITTNER

Division of Cancer Biology, Department of Physiology, University of Minnesota Medical School, Minneapolis, Minnesota

The development of mammary tumors in mice is usually dependent upon the action of three factors, *viz.*, the inherited susceptibility, hormonal stimuli, and the agent normally transferred in the mother's milk.[12] A small percentage of mammary tumors may, however, develop in animals in which one or more of the causative factors cannot be demonstrated with our present methods of assay. Another genetic factor has been demonstrated to play a part in the genesis of mammary cancer in virgin females. This is mediated through the genic control of the hormonal mechanism and has been called the inherited hormonal influence.[26, 41, 25] Some data are available on the physiological effects of this factor[29, 42-44, 48-50] and its genetic relationship to the inherited susceptibility for mammary cancer in some stocks.[25]

The mammary tumor milk agent was demonstrated by fostering the young born to mothers of cancerous strains by females of low cancerous stocks.[8] If the fostered mice did not obtain the agent from the milk before they were transferred, they and their descendants usually showed a low incidence of mammary cancer. Fostering after the mothers of cancerous strains had nursed their young for twenty-four hours or longer had little effect in reducing the incidence,[9, 4, 10] and the females did not have to be cancerous to transfer the agent.[8] It was also found that the agent was present in the milk secreted after the first day.[9, 4] The removal of the young from the uteri of females of cancerous strains was found to be more effective than foster nursing in preventing the development of mammary cancer in mice.[1]

By feeding or injecting extracts of tissues from mice with the agent into suitable test animals (descendants of fostered susceptible mice which have been observed for several generations, or hybrids which had maternal parents from stocks without the agent), it has been possible to study the distribution of the agent and its characteristics. It has been demonstrated to be present and active, based on the development of mammary cancer in the test animals, in the following tissues: thymus,[9, 11] spleen,[9, 11, 5, 30] lactating mammary tissue,[13, 3] spontaneous mammary cancer,[15, 7, 3] Harderian gland,[27] liver (unpublished),* and whole blood.[33] Although some workers[36] have noted few, if any, mammary tumors to develop in mice injected with either whole blood or the blood fractions, we have found, in some studies,[19, 20] that the agent was active in both suspensions of blood cells and the serum. By adding red blood cells from mice without the agent, it has been possible to adsorb the agent from the serum of cancerous animals.[23]

---

* Assisted by grants from the Citizens Aid Society of Minneapolis, the Graduate School Cancer Research Fund of the University of Minnesota, the Jane Coffin Childs Memorial Fund for Medical Research, the Floyd B. Olson Memorial Fund, the Elsa U. Pardee Foundation, the Minnesota Cancer Society, and the American Cancer Society.

The agent will remain active following lyophilization,[17] filtration through Seitz[16] and Berkefeld filters,[17, 3] treatment with glycerin,[16, 3] and desiccation.[31, 21, 32-34] By the injection of cell-free filtrates into the developing chick embryos, it has been found that the agent may be recovered from either filtered or unfiltered yolk after one passage[24] and from unfiltered yolk after nine serial passages in eggs (unpublished). The agent will also survive for at least ten serial passages of mammary cancer in mice which did not have it.[24] It becomes inactive at pasteurization temperature within thirty minutes.[3, 7]

Several investigators[1, 17, 5, 2, 21, 32] have shown that animals of susceptible stocks become more resistant to the agent with increasing age. However, this resistance may be overcome by repeated injections.[32] Although the older mice may not give rise to mammary tumors, they may transfer the agent to their progeny.[14, 18]

In earlier studies, it was stated that the amount of the agent which an animal received, determined the incidence and average cancer age. However, recently it has been found that this may not be the case.[20] Extracts of equivalent amounts of tissues may produce in mice of some ages higher incidences when diluted 10^{-4} than they do at 10^{-1}. In younger animals, the difference may not be as evident. In some groups, the prevention of nursing will influence the incidence, while in others it will not.[22, 23] The agent may remain active in dilutions of 1:1,000,000, and approximately 25 per cent of the injected mice may become cancerous.[6, 22, 23]

In fractions which had been diluted 10^{-2} to 10^{-6}, it was noted that a larger percentage of the test animals developed mammary tumors after they had been injected with extracts of lactating mammary tissue than with either spontaneous mammary cancer or transplanted mammary cancer from the same stock.[6] When mice of a different genetic constitution were employed, there was no difference in the activity of the extracts of mammary and spontaneous tumor tissue, but both produced a higher incidence than did either suspended blood cells or serum from cancerous mice of the same stock.[23]

The injection of extracts of spontaneous mammary cancer from mice stimulated the formation of antibodies in rabbits[3, 38-40] and rats.[38-40] It

---

In earlier experiments,[13] we were unable to obtain mammary cancer in mice which had been given extracts of liver from mice with the agent.

has been found that the serum from these immunized animals would neutralize the agent *in vitro*. If the anti-cancer serum was given to mice before they received the agent, they would not develop mammary cancer[3] but, if the mice obtained the agent before the anti-serum, no effect has been seen in either the injected mice or their progeny (unpublished). The immune serum has a cytotoxic effect on mammary cancer cells *in vitro*.[37]

In centrifugation studies, it has been observed that the agent was either sedimented or associated with particles of the cell which were thrown down in gravitational fields of 20,000,[28] 60,000,[45] and 110,000.[52] In another experiment, essentially all of the agent was sedimented at 18,000x g. for a period of one hour, the greatest activity being found in the fractions with the larger particles or the microsomes at dilutions of 10^{-2} to 10^{-6}.[6]

It was suggested that the presence of the milk agent might influence the formation of lateral buds along the ducts of the mammary glands of mice of some stocks.[51, 47] However, these observations could not be confirmed by the use of fostered and unfostered mice of other cancerous stocks.[12] Estrogenation may produce more arborization in young mice when they have the agent, but these differences may not be evident at later ages.[46] A possible explanation[35] for these different observations has been suggested as due to strain differences in the architecture of the normal mammary gland.

The characteristics of the mammary tumor milk agent, its small size, ability to propagate in the living cell, and its antigenic properties, would classify it as an infectious agent or virus. Its exact role in the genesis of mammary cancer has not been determined, and there is no evidence that it may be active in the development of any other type of cancer in mice.

LITERATURE CITED

1. **Andervont, H. B.**
   1941. J. Nat. Cancer Inst. **2**: 13–16.
2. **Andervont, H. B.**
   1945. J. Nat. Cancer Inst. **5**: 397–402.
3. **Andervont, H. B., & W. R. Bryan**
   1944. J. Nat. Cancer Inst. **5**: 143–149.
4. **Andervont, H. B., & W. J. McEleney**
   1939. Pub. Health Rep. **54**: 1597–1603.
5. **Andervont, H. B., M. B. Shimkin, & W. R. Bryan**
   1942. J. Nat. Cancer Inst. **3**: 309–318.
6. **Barnum, C. P., Z. B. Ball, & J. J. Bittner**
   1946. Cancer Research **6**: 499.
7. **Barnum, C. P., Z. B. Ball, J. J. Bittner, & M. B. Visscher**
   1944. Science **100**: 575–576.
8. **Bittner, J. J.**
   1936. Science **84**: 162.

9. **Bittner, J. J.**
1939. Am. J. Cancer **35**: 90–97.
10. **Bittner, J. J.**
1939. Pub. Health Rep. **54**: 1642–1650.
11. **Bittner, J. J.**
1939. Pub. Health Rep. **54**: 1827–1831.
12. **Bittner, J. J.**
1939. Pub. Health Rep. **54**: 1590–1597.
13. **Bittner, J. J.**
1940. Proc. Soc. Exp. Biol. & Med. **45**: 805–810.
14. **Bittner, J. J.**
1941. Cancer Research **1**: 113–114.
15. **Bittner, J. J.**
1941. Science **93**: 527–528.
16. **Bittner, J. J.**
1942. Science **95**: 462–463.
17. **Bittner, J. J.**
1942. Cancer Research **2**: 710–721.
18. **Bittner, J. J.**
1943. Cancer Research **3**: 441–447.
19. **Bittner, J. J.**
1944. Bull. Minn. Med. Foundation **4**: 94–96.
20. **Bittner, J. J.**
1945. Proc. Soc. Exp. Biol. & Med. **59**: 43–44.
21. **Bittner, J. J.**
1945. A. A. A. S. Research Conference on Cancer : 63–95.
22. **Bittner, J. J.**
1946. Cancer Research **6**: 493.
23. **Bittner, J. J.**
1946. Univ. Buffalo Centennial Program. In press.
24. **Bittner, J. J., C. A. Evans, & R. G. Green**
1945. Science **101**: 95–97.
25. **Bittner, J. J., & R. A. Huseby**
1946. Cancer Research **6**: 235–239.
26. **Bittner, J. J., R. A. Huseby, M. B. Visscher, Z. B. Ball, & F. W. Smith**
1944. Science **99**: 83–85.
27. **Bittner. J. J., & C. J. Watson**
1946. Cancer Research **6**: 337–343.
28. **Bryan, W. R., H. Kahler, M. B. Shimkin, & H. B. Andervont**
1942. J. Nat. Cancer Inst. **2**: 451–456.
29. **Deringer, M. K., W. E. Heston, & H. B. Andervont**
1945. J. Nat. Cancer Inst. **5**: 403–406.
30. **Dmochowski, L.**
1944. Brit. J. Exp. Pathol. **25**: 119–120.
31. **Dmochowski, L.**
1944. Brit. J. Exp. Pathol. **25**: 138–140.
32. **Dmochowski, L.**
1945. Brit. J. Exp. Pathol. **26**: 192–197.
33. **Dmochowski, L.**
1945. Brit. J. Exp. Pathol. **26**: 267–269.
34. **Dmochowski, L.**
1945–1946. Imperial Cancer Research Fund 43rd Annual Report : 9.
35. **Gardner, W. U.**
1945. A. A. A. S., Research Conference on Cancer : 95.

36. **Graff, S., H. T. Randall, G. E. Carpenter, & C. D. Haagensen**
1946. Science **104**: 289.
37. **Green, R. G.**
1946. Proc. Soc. Exp. Biol. & Med. **61**: 113–114.
38. **Green, R. G., & J. J. Bittner**
1946. Cancer Research **6**: 499.
39. **Green, R. G., M. M. Moosey, & J. J. Bittner**
1945. Cancer Research **5**: 588–589.
40. **Green, R. G., M. M. Moosey, & J. J. Bittner**
1946. Proc. Soc. Exp. Biol. & Med. **61**: 115–117.
41. **Heston, W. E., & H. B. Andervont**
1944. J. Nat. Cancer Inst. **4**: 403–407.
42. **Huseby, R. A., & J. J. Bittner**
1946. Cancer Research **6**: 240–255.
43. **Huseby, R. A., F. W. Smith, & J. J. Bittner**
1946. Cancer Research **6**: 494.
44. **Huseby, R. A., F. W. Smith, & J. J. Bittner**
1946. Cancer Research **6**: 494.
45. **Kahler, H., & W. R. Bryan**
1943. J. Nat. Cancer Inst. **4**: 37–45.
46. **Shimkin, M. B.**
1945. A. A. A. S., Mammary Tumors in Mice : 85–122.
47. **Shimkin, M. B., H. G. Grady, & H. B. Andervont**
1941. J. Nat. Cancer Inst. **2**: 65–80.
48. **Smith, F. W.**
1945. Science **101**: 279–281.
49. **Smith, F. W.**
1946. Cancer Research **6**: 494.
50. **Smith, F. W., & J. J. Bittner**
1945. Cancer Research **5**: 588.
51. **van Gulik, P. J., & R. Korteweg**
1940. Proc. Nederl. Akad. Wetenschappen **43**: 891–900.
52. **Visscher, M. B., R. G. Green, J. J. Bittner, Z. B. Ball, & H. A. Siedentopf**
1942. Proc. Soc. Exp. Biol & Med. **49**: 94–96.
53. **Wooley, G. W., L. W. Law, & C. C. Little**
1941. Cancer Research **1**: 955–956.

DISCUSSION OF THE PAPER

Dr. Arthur Shapiro *(Brooklyn, N. Y.)*:

If the rabbit anti-mouse tumor serum produces specific cytolysis of mouse tumor cells through the action of an antibody against a tumor "virus", it seems likely that this reaction could be used for the assay of "virus" in a manner analogous to the titration of diphtheria toxin by antitoxin. If it were determined in advance that a particular dilution of a standard "virus" suspension was just able to neutralize the cytolytic effect, then either "virus" or serum (whichever is more stable) could be used as a standard of reference for the *in vitro* determination of the activity of other preparations with which it could be compared.

SOME BASIC CONTRIBUTIONS TO THE CANCER PROBLEM FROM THE STUDY OF PLANTS

By WILLIAM J. ROBBINS

Department of Botany, Columbia University, and The New York Botanical Garden, New York, N. Y

I must preface my remarks on this subject by emphasizing that cancer is a disease of animals, and that we must be conservative in applying information gained from the study of overgrowths in plants to similar phenomena in animals. Overgrowths developing spontaneously or induced by microorganisms, by insects, by some types of grafts, by crossing certain plant species, by viruses, by wounds, and by various physical or chemical treatments occur in plants. They are variously known as mutations, calluses, galls, tumors, fasciations, nodules, or even plant cancers. Some of them show analogies to cancer in animals, but it is always dangerous to argue from analogy, and I wish to make clear that I appreciate the inherent weaknesses in an attempt to apply observations made on one organism to another, especially when they are so distantly related as plants are to man.

Nevertheless, I think we can agree that cancer is fundamentally a growth problem, that one of its basic aspects is the transformation of cells from a more or less static into a rapidly growing condition. I believe we can agree, further, that the processes which occur in the growth of plants and animals are basically alike. I need not remind the reader that at least three vitamins (biotin, pantothenic acid, and folic acid) were first discovered by observing their importance for the growth of bacteria or yeasts, and that a great deal has been learned of the intermediary metabolism of carbohydrates in the human body through a study of the fermentation of sugar by yeast. Fundamental information on the growth of plants may easily be of importance to the cancer problem, particularly if the phenomena studied are those concerned with spontaneous or induced changes resulting in increased rates of growth, especially when those changes are persistent and result in more or less disorganized growth. Any instance of this sort which occurs in plants deserves careful study and analysis. An elucidation of any one of them may throw light on abnormal growth in animals. A few such instances in plants will be briefly discussed here.

Those spontaneous mutations which occur in the mycelium of many species of fungi deserve more consideration, I believe, than they have hitherto received. They have been variously known as saltations, sectorial mutations, or pleomorphisms.

We have studied one of the dermatophytes, *Trichophyton mentagrophytes* or *Trichophyton gypseum,* which develops vegetative mutations to a high degree.[7, 8] This fungus. pathogenic for man and other animals, is, as a rule, a rather slow-growing and freely sporulating organism when isolated from the host. We have called this the normal form.

As cultures of the normal form of this organism age, portions of the mycelium change in growth habit: they become white, fluffy, more or less sterile, and much more vigorous than the original mycelium. Such forms have long been known, in the dermatophytes, as pleomorphic forms. They can be isolated and grown in pure culture where they retain their morphological characters and ability to grow rapidly. In culture, pleomorphic forms overgrow the normal form and, unless suitable precautions are taken, the normal form of *T. mentagrophytes* will disappear and only the rapidly growing pleomorphic forms will remain.

The normal form can be maintained indefinitely if transfers are made to fresh media at weekly intervals. We have kept it in culture for more than 3 years by this procedure. Pleomorphic forms, too, can be kept in culture apparently unchanged if transferred at weekly intervals.

The spontaneous change from normal to pleomorphic occasionally appears in cultures incubated at 35° C. as early as 8 days after inoculation, though two or three weeks are usually required, and within 4 or 5 weeks, at most, all cultures will have become pleomorphic. It is not, therefore, a hit-or-miss process. Merely by allowing the cultures to age, all of them will change from a slow-growing to a rapidly growing form.

Many pleomorphic types can be isolated which differ in rate of growth, degree of sterility, morphological character of the colony, in pigment production, or in some physiological character. All older cultures, however, are characterized by the very vigorous, white, fluffy, almost sterile form. This may be because it overgrows the less vigorous types, or, perhaps, because the less vigorous types mutate, in time, toward the fast-growing forms.

The growth of the most vigorous forms is many times that of the normal. In a liquid medium containing asparagine as a nitrogen source, the normal form produces 3 or 4 mg. of dry matter in three weeks. A pleomorphic form, in the same time in the same medium, may form 100 mg. or more.

Here we have, in an organism which can be readily grown in a test tube, an example of spontaneous change from a slow-growing to a rapid-growing condition which shows some analogies to the changes which occur when more or less static cells in the animal body are transformed into cancer cells. If we could answer the problems on growth presented by this phenomenon, such answers might elucidate some aspect of cancer.

What causes pleomorphisms to occur? There appear to be three possibilities. They may result from the effect, on the mycelium, of the accumulation of by-products of fungus metabolism in the medium. Dr. Jacques Duché tells me that he was able to keep *Trichophyton gypseum* in a normal condition, without the development of pleomorphisms, for a period of 2½ months, by growing the fungus under conditions such that it was continuously bathed in fresh culture medium. These observations would support the idea that the spontaneous mutation from a slow-growing condition to a rapid one, in this fungus, is induced by the accumulation of metabolic products in the culture medium. If this is correct, then the cells of the fungus mycelium must differ in susceptibility to these products, because the entire mycelium does not become pleomorphic and only small scattered sectors, perhaps individual cells, are affected. For example, on an agar medium in a Petri dish 18 cm. in diameter which is completely covered with the mycelium of the normal fungus, a few scattered spots of a pleomorphic form appear. The whole mass of mycelium does not become pleomorphic (PLATE 14, FIGURE 1).

Another possibility is that the pleomorphism is induced by an exhaustion of one or more constituents of the medium. This does not seem likely in view of the richness and variety of media on which pleomorphisms develop.

The third possibility is that the mutation is associated, in some way, with the age of individual hyphae. It cannot be associated with the age of the entire mycelium, because we have never observed the development of pleomorphisms on the young growing hyphae but always on the older parts of a colony where the age of a hypha, the accumulation of by-products, or the exhaustion of some essential substance might be possible causal factors. As long as the mycelium grows over fresh medium, it remains normal. At least, we have observed this to occur over a period of more than 60 days (PLATE 14, FIGURE 2).

Can the development of pleomorphisms be prevented? We have not been successful in any attempt of this kind, except by transfers, at weekly intervals, to fresh media, as previously mentioned. Nor have we had any evidence that the pleomorphic form can be induced to revert to the normal form. Various modifications of the nutrient medium, as well as a number of growth inhibitors, both natural and synthetic, have been used, without preventing the formation of pleomorphic forms or inducing them to revert once they have developed.

Why do the pleomorphic forms grow so much more rapidly than the normal? We believe that this is due to the fact that the mutant has developed a more effective mechanism for transforming compounds of nitrogen into cell substance. This mechanism is probably new enzyme systems or larger amounts of some which already exist in the normal form. The evidence for this assumption is indirect and is as follows:

The normal form of *Trichophyton mentagrophytes* grows with asparagine, or any one of 14 amino acids, as a source of nitrogen, but is unable to utilize ammonia effectively. It appears that it transforms asparagine, or any one of the several amino acids, without previous ammonification into all the various amino acids required for the construction of its protoplasmic proteins. There are, however, no indispensable amino acids in the sense that the fungus fails to grow unless a particular amino acid is furnished in the nutrient medium. Although the normal form of *T. mentagrophytes* is capable of making from a single amino acid all those needed for its cell substance, it grows better with a mixture of these. We have interpreted this to mean that the amino acids in a mixture are incorporated into fungus protein more rapidly than they are supplied by the metabolic transformation of a single one.

The rapidly growing pleomorphic forms, on the other hand, grow fairly well with ammonium salts as the sole source of nitrogen. This difference in the relation of the normal and the pleomorphic forms to ammonium salts we have interpreted as indicating that, in changing from a slow-growing to a rapidly growing condition, the fungus has developed a better mechanism for utilizing inorganic nitrogen.

By supplying the normal form with a mixture of amino acids, as much dry matter can be obtained as the pleomorphic form produces with a single amino acid. This also suggests that, in becoming pleomorphic, the fungus has developed a more effective means of transforming a single organic nitrogen source into cell substance.

Are the pleomorphic forms genic or cytoplasmic mutations? This question cannot be tested by ordinary genetical methods, because the sexual form of *Trichophyton mentagrophytes,* if it exists, has not been discovered.

Another group of investigations of particular interest to the general subject under discussion, are those concerned with the development, in plants, of tumor tissue which is capable of indefinite growth in a relatively simple medium and of autonomous growth when grafted into the normal plant. These include the bacteria-free tumors resulting from infection by the bacterium, *Phytomonas tumefaciens,* the tumors induced by grafting bacteria-free tumor tissue on normal tissue reported by de Ropp, those occurring on certain types of species hybrids of *Nicotiana,* and the wound-virus tumors described by Black.

The overgrowths caused by *Ph. tumefaciens,* generally referred to as crown gall, have been studied for many years. Irwin F. Smith emphasized the similarity between crown gall of plants and cancer of animals. Recent observations have added a great deal to our knowledge of these interesting growth abnormalities but still leave many fundamental questions unanswered. Jensen[5] found that the tumorous condition of the host cells of beets and mangels induced by *Ph. tumefaciens* persists after the bacteria have disappeared. According to Jensen, the bacteria died in the older tumors caused by *Ph. tumefaciens*, and careful investigation of a large number of spontaneous tumors from mangels and sugar beets, taken up in the autumn, led only in a single instance to the isolation of the bacterium. Tissue from these spontaneous beet tumors was easily transplanted to normal roots and produced fresh tumors which were due exclusively to the growth of the transplanted tissue. However, Jensen says that, while tumor transplantation proved effective for three or four generations, the fourth generation of implantation succeeded, but the abnormal proliferative power of the tissue had disappeared. He concluded that the cells of the tissue, under the influence of the bacterium, become altered for a series of cell-generations and develop the abnormally increased proliferative power of tumor cells proper, which is independent of continued stimulation by the bacteria.

Jensen's observations were confirmed and extended, more recently, by White, Braun, and others.[11] They demonstrated that the bacteria-free tumor tissue can be grown indefinitely on a simple medium and develops autonomously as tumor tissue when grafted on the normal plant. The bacteria cause the change from normal to tumorous condition in a short time. Tumors develop in *Vinca rosea* after but 36 to 48 hours' exposure to living bacteria, and the full change to tumor tissue occurs after 4 days.

Amongst many observations, too numerous to detail here, was the demonstration, by Braun and Laskaris,[13] that the combined effect of infection by an avirulent strain of *Ph. tumefaciens* and treatment with growth-promoting substances such as indole-acetic acid was necessary for tumor production on the tomato. Neither was effective alone. This indicates that two factors, at least, are concerned in the production of tumors by *Ph. tumefaciens.*

De Ropp[9] found that the bacteria-free tumor tissue grafted *in vitro* on pieces of normal stem 0.5 cm. in length induced secondary tumors in about 50 per cent of the successful grafts. These secondary tumors can be grown *in vitro* and evidence autonomous growth when grafted into normal plants. If the graft of tumor and normal tissue did not take, or if the tumor tissue was separated from the normal tissue by a thin layer of agar, no secondary tumors were observed. These observations demonstrate that tumefaction may be induced by bacteria-free tumor tissue as well as by the bacteria themselves. It also suggests that some influence from the bacteria-free tumor tissue is transmitted through living tissue at least for short distances and induces the change from normal to tumor tissue. Whether this influence is chemical, virus-like, or a cytoplasmic entity, has not been determined.

Black's observations[1, 2] on the wound-tumor virus demonstrate another method of tumor induction in plants. This virus, which affects at least 43 species of plants in 20 families, is systemic. The virus produces tumors on the roots of many kinds of plants and on the stems of some others. Wounds are involved in the origin of at least some of the tumors on the stems of infected plants and, since lateral roots wound the overlying tissues in emerging from their point of origin in the pericycle, there seems to be an interrelation between the presence of the virus and injury in the production of tumors on roots. The tumor tissue can be grown indefinitely *in vitro* and can be successfully grafted into normal plants. It retains the virus, which becomes systemic in a host on which a bit of tumor tissue is grafted.

Still another kind of origin of autonomous tumors in plants results from hybridizing species of *Nicotiana.* Kostoff, in 1930, reported the occurrence of abundant tumors on plants of the hybrids of certain *Nicotiana* species. These were later (1934) studied by Whitaker.[10] The tumor tissue was cultivated *in vitro* by White and demonstrated to produce tumors when grafted on normal plants. These tumors do not appear to be parasitic in origin; no causal agent can be transmitted from tumor to healthy *Nicotiana* plants; and the causal factor is not transmitted from hybrid to parent over a graft union. The tumors apparently result from the hybridity. Whitaker concluded, from an analysis of available data, that hybrids with tumors are formed when representatives of the *Nicotiana alata* group with 9 chromosomes are used as the male parent and other Nicotianas which have 24 chromosomes are the female parents, but that no such formation takes place when the reverse cross is made. All the available information, according to Whitaker, indicates that tumor production in hybrids of *N. Langsdorffii* and *N. glauca* is the result of cytoplasmic disturbance occasioned by the introduction of the chromosome complement of *N. Langsdorffii* (used as the pollen parent) into the cytoplasm of *N. glauca* (used as the seed parent).

From this brief survey, it is clear that plant tumors characterized by more or less unorganized and uncontrolled growth, capable of indefinite growth *in vitro* on simple media and of autogenous growth *in vitro,* are induced by a considerable variety of causative agents. They can be induced by inoculation with a bacterium, *Phytomonas tumefaciens,* by bacteria-free tumor tissue, by the combined effect of an attenuated strain of *Ph. tumefaciens* and a growth-promoting chemical, by infection with a specific virus plus wounding, and as the result of certain species crosses where a virus, wounding, pathogenic organisms, or the application of chemicals are not involved.

There are many other kinds of overgrowths in plants which either do not possess the ability of the tumors I have briefly described to grow indefinitely *in vitro* or *in vivo,* or have not been demonstrated to have that ability. Club root of cabbage and other crucifers results from infection by *Plasmodiophora brassicae,* a myxomycete. Insect galls are common on plants. These are characterized by an abnormal but specific growth pattern superimposed on the normal tissue by a foreign living organism. Calluses grow from the cut ends of stems or from wounds. By treatment with chemicals, as, for example, indole-acetic acid, tumors in plants have been induced which produce massive growths. Fasciations, spontaneous or induced, are another class of abnormal overgrowths common in plants.

In spite of the progress which has been made in our knowledge of autonomous tumors and other types of abnormal plant growths, the basic and fundamental problems remain. Why and how does an actively growing embryonic or meristematic cell change into a relatively static, mature, differentiated condition? On the contrary, why and how does the change of non-growing cells into those capable of active growth occur, or that of cells which grow normally into tumor cells?

We know little of the basic factors involved in these changes. Our ignorance is sometimes concealed by such words or phrases as growth regulators, organizers, the whole being greater than the sum of its parts, entelechy, primitive condition, and so on. But why does a section of a sunflower stem which lacks a bud but contains embryonic cells in its cambium, fail to grow to any extent when grafted into the normal stem where tumors induced by infection with *Phytomonas tumefaciens* grow well?

One explanation which may be offered is that inhibitors from the entire plant limit the development of the transplant of the normal tissue but are ineffective on the tumor tissue. It is not difficult to cite examples in the plant kingdom where inhibitors produced by an organism affect its growth. Among the microorganisms, the development of specific substances which inhibit the organism's growth are well known, as, for example, lactic acid accumulating in cultures of lactic acid bacteria, alcohol for the yeasts, or acetic acid for the acetic acid bacteria. Fawcett demonstrated, some years ago, that filamentous fungi grow at an unchanging rate if constantly supplied with fresh media. The growth of dormant axillary buds when the tip of the plant stem is cut off, has been explained by the removal of inhibitors formed by the stem tip.

It does not seem probable, however, that inhibitors from the entire plant can account for the failure of normal plant tissue to grow when grafted into the stem, because the bit of normal tissue removed from the influence of the plant evidences little growth *in vitro* on a medium on which the tumor tissue grows indefinitely. How can the difference in growth power evidenced *in vitro* between the normal stem tissue and the tumor tissue derived from it be accounted for? It is evident that the tumor tissue on a simple medium carries out all the series of chemical reactions which result in the construction of new cell substance. The normal tissue is not able to do this.* Is this because the cell substance of the tumor tissue is simpler, more easily synthesized, and requires a less elaborate mechanism for its construction? Is it because normal tissue develops intracellular inhibitors which limits its growth? Or is it because, in becoming tumorous, the cells develop a more effective mechanism for making new protoplasm, or a combination of these?

---

* Gautheret and Nobecourt have been able to grow normal stem tissue *in vitro* by adding small quantities of indole-acetic acid to the medium. It is not known how the indole-acetic acid acts in inducing the tissue to grow. For references to literature, see White.[12]

It may be pointed out that some kinds of plant tumor tissue do not form chloroplasts or differentiate roots, stems, and leaves, which might be interpreted to mean that the cell substance of the tumor tissue is less complicated than that of normal tissue. At the same time, we cannot exclude the possibility that the normal cell makes intracellular inhibitors which the tumor tissue does not. There are, also, numerous examples among the microorganisms where the rate of growth is determined by the ability of the organism to synthesize one or more vitamins, amino acids, or some other essential metabolite.

Greenstein[4] has said, "The history of cancer research is such as to effectively discourage any excursion into a dogmatic universalism." The same caution must be used in searching for the causes for the differences in growth power between normal and tumorous tissue, between meristematic or embryonic differentiated mature cells.

To determine whether these changes are genic or cytoplasmic, and whether they are associated with morphological entities or not, is important. It is equally important, however, to learn, if we can, the internal physiological factors which limit the growth of cells,[6] because it is the modification of these factors which change a normal meristematic cell or a static and mature one into those which we call tumorous. What happens when mature cells regain the power of making new cell substance? Why, when this occurs, do such embryonic cells not follow the growth pattern of normal cells and after a limited period mature once more, instead of growing for an unlimited period and in a more or less disorganized fashion? Do cells mature and cease to grow because of the accumulation of inhibitors, because they lose some essential part of their metabolic machinery (say, an enzyme system), or for some other reason? Do mature cells return to a growing condition spontaneously or under the action of some extraneous physical, chemical, or biological agent; because of the neutralization or elimination of growth inhibitors; because a new synthetic mechanism has been formed by which cell substance perhaps of simpler and less complicated constitution is more quickly constructed; because parts of the machinery necessary for the synthesis of new protoplasm, lost in the process of maturation, are once more reestablished; or for some other reason? These are fundamental questions for which I have no categorical answers. I believe, however, that answers may be sought for them by using plants as well as animals as experimental material and, perhaps, in some respects, plant material may furnish the shorter road to the solutions in which we are all interested.

LITERATURE CITED

1. **Black, L. M.**
   1945. A virus tumor disease of plants. Am. J. Bot. **32**: 408–415.
2. **Black, L. M.**
   1946. Plant tumours induced by the combined action of wounds and virus Nature **158**: 56.
3. **Braun, A. C., & T. Laskaris**
   1942. Tumor formation by attenuated crown-gall bacteria in the presence of growth-promoting substances. Proc. Nat. Acad. Sci. **28**: 468–477.
4. **Greenstein, Jesse P.**
   1945. Enzymes in normal and neoplastic animal tissues. Am. Ass. Adv. Sci. Research Conference on Cancer: 192–222.
5. **Jensen, C. O.**
   1918. Investigations upon tumor-like formations in plants. (English summary) D. K. Veterinaer. og Landboliojskole. Aarsskrift: 91–143.
6. **Robbins, William J.**
   1942–43. Internal factors limiting growth in plants. The Harvey Lectures **38**: 187–198.
7. **Robbins, William J., & Roberta Ma**
   1945. Growth factors for *Trichophyton mentagrophytes*. Am. J. Bot. **32**: 509–523.
8. **Robbins, William J., & Ilda McVeigh**
   1946. Effect of hydroxyproline on *Trichophyton mentagrophytes* and other fungi. Am. J. Bot. **33**: 638–647.
9. **de Ropp, R. S.**
   1947. The growth-promoting and tumefacient factors of bacteria-free crown-gall tumor tissue. Am. J. Bot. **34**: 248–261.
10. **Whitaker, T. H.**
    1934. The occurrence of tumors on certain *Nicotiana* hybrids. J. Arnold Arboretum **15**: 144–152.
11. **White, P. R., & A. C. Braun**
    1942. A cancerous neoplasm of plants. Autonomous bacteria-free crown-gall tissue. Cancer Research **2**: 597–617.
12. **White, P. R.**
    1946. Plant Tissue Cultures. II. Bot. Review **12**: 521–529.

DISCUSSION OF THE PAPER

Dr. Arthur Shapiro (*Brooklyn, N. Y.*):

The "pleomorphic" variant described by Dr. Robbins presents interesting analogies with the behavior of neoplasms in tissues. However, it is not clear from his data whether he has taken steps to distinguish between factors which favor the occurrence of mutation and factors which favor the detection of mutation already present.

It would appear that, since the pleomorphic forms grow faster than the normal form, they should be readily detected whenever they occur. However, the experiments in the "racing tube" seem to indicate that the rate of spread of the normal culture is about equal to that of the pleomorphic culture.

In the evaluation of the likelihood of detection of the mutant under given cultural conditions, this biological competitive growth rate would seem to be more important than the growth rate measured as mass of tissue formed per unit time.

It should also be mentioned that, in bacterial cultures, it is common to find mutants arising (in aged cultures) which outgrow the normal type in the aged media but are definitely outgrown by the normal type in fresh media.

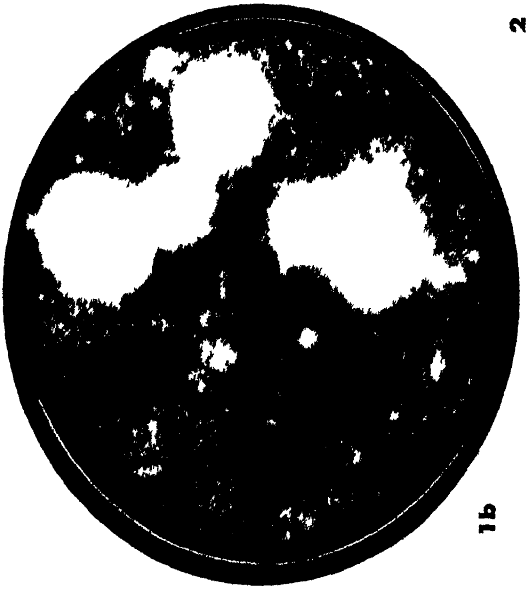
In general, then, it is important to recognize, distinguish, and measure growth rates and mutation rates separately, if one wishes to draw valid conclusions from experiments on large populations of microorganisms.

PLATE 14

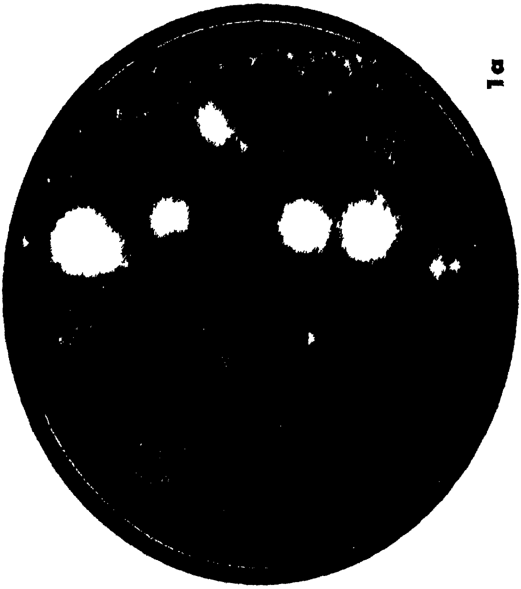
PLATE 14

FIGURE 1. A, *Trichophyton mentagrophytes*, normal form, 17 days old, on peptone-dextrose agar. The development of five pleomorphic forms is obvious B, same dish 3 days later. Note rapid growth of pleomorphic forms.

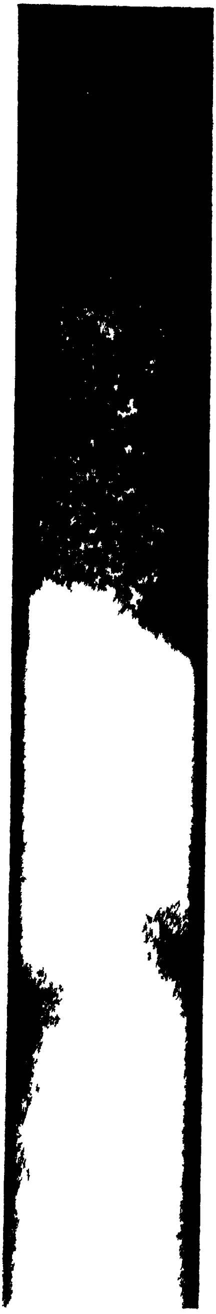
FIGURE 2. *Trichophyton mentagrophytes*, after 59 days' growth on dextrose-asparagine agar medium supplemented with casein hydrolysate. Horizontal tube partially filled with medium was inoculated with the normal form at left. Note that pleomorphisms have developed on older portion, but that the young mycelium advancing over fresh medium remains normal, after 59 days' growth.



1b



1a



2

CHEMICALLY INDUCED MUTATIONS AND THEIR BEARING ON CARCINOGENESIS

By E. L. TATUM

Yale University, New Haven, Connecticut

Increasingly vigorous attempts have been made, during the past decade, to analyze the complex problem of carcinogenesis in terms of the mutational, viral, and other theories of its causation (see Furth, Boon, and Kaliss, 1944). It is difficult, perhaps unnecessary, to reconcile these apparently opposed views, or to correlate them with the proven roles of carcinogenic chemicals and irradiation in the production of experimental cancer. Nevertheless, during the past few years, developments in this and other fields have brought new evidence regarding the relation of genes to enzymes and biochemical reactions, and regarding the relation of gene to cytoplasm, which may modify our older concept of mutation. Finally, recent evidence of the production of mutations by chemical treatment, and the examination of the mutagenic action of at least one carcinogen, methylcholanthrene, may bear directly on the problem. Even if the various theories cannot as yet be satisfactorily reconciled, a consideration of possible relationships may not be amiss and may suggest new approaches to carcinogenesis.

The theory that cancer is due to somatic mutations is based mainly on the following points, the first three of which constitute only inferential support:

(1) That susceptibility and resistance to carcinogenesis are determined, in experimental animals, by hereditary factors, susceptibility ordinarily being dominant to resistance.

(2) That transplantability of normal tissues seems to depend on the genetic and antigenic constitutions of both host and transplant.

(3) That tumor cells may differ from the normal tissue from which they are derived in transplantability and, hence, may differ also in genetic constitution. The mutations indicated by this are not necessarily related to those mutations postulated in carcinogenesis.

(4) That the change from a normal to a malignant cell may occur at random in small groups of cells or even in single cells, is a potential reaction of all cell types, and may involve heritable changes, perhaps of a sudden abrupt nature.

(5) That the effects of x-ray and ultraviolet irradiation in carcinogenesis may be correlated with the well-established mutational effects of these agents in plants, animals, and microorganisms.

The many biochemical and enzymatic changes correlated with the development of malignancy (Potter, 1944; Burk and Winzler, 1944; Greenstein, 1945) could result from a single primary genetic or enzymatic change and, hence, are not in disagreement with the somatic mutation theory.

One of the principal difficulties in the acceptance of this theory has been the inability to account satisfactorily for the effects of chemical agents such as the carcinogens. Many attempts to detect mutational effects of chemicals have been inconclusive, because of the unsuitability of the biological material for genetic analysis, and have no direct bearing on cancer since the agents used are not carcinogenic. To a large extent, this is true for the investigations of the effects of colchicine on a variety of cells, and of camphor (Bauch, 1941) and acenaphthene (Levan and Sandwell, 1943) on yeast. Recent investigation by Levan (1946) has given cytological support to earlier assumptions of the effect of these agents on chromosome-doubling in yeast and has added butyl alcohol and benzene to the agents inducing mitotic disturbances in this organism. The absence of direct genetic or cytological evidence makes it difficult to evaluate any postulated mutational effect of cyanide on yeast (Stier and Castor, 1941) and of nitrite and other agents on *Aspergillus* (Steinberg and Thom, 1940).

Other attempts to demonstrate mutagenic action of chemicals with genetically suitable material have not been conclusive, perhaps because of the relative ineffectiveness of the agents used, or because of difficulties in establishing effective concentrations in germ-plasm. This is true of most investigations with *Drosophila,* although possibly significant effects have been obtained here with copper sulfate (Law, 1938), phenol (Hadorn and Niggli, 1946), and with sulfonamides (Thomas and Chevais, 1943). Finally, it is as yet impossible to detect and demonstrate mutations in tissues of higher animals by the direct application of classical genetic techniques. The recent demonstrations, by Auerbach and Robson, of the effects of mustard gas in inducing germinal as well as somatic mutations in *Drosophila* (Auerbach and Robson, 1946; Gilman and Philips, 1946; Dixon and Needham, 1946) may lead to a dissolution of this apparent impasse.

Examination of the effects of this chemical, $\beta\beta'$-dichlorodiethylsulfide (Horowitz *et al.,* 1946) and of the related nitrogen compound, $\beta\beta'$-dichlorodiethylmethylamine (Tatum, unpublished results) on *Neurospora,* has proved that both agents produce gene mutations apparently analogous, in effect, to those produced by irradiation (TABLE 1). It may be more than coincidence that some of the N-mustard mutations in *Neurospora* are genetically, and probably biochemically, different from mutations so far obtained with irradiation. The ability of these chemicals to induce mutations is further substantiated by the effect of N-mustard

TABLE 1

MUTATIONS PRODUCED IN *Neurospora* BY TREATMENT WITH SULFUR MUSTARD AND NITROGEN MUSTARD*

	Total cultures tested	Morphological mutants		Biochemical mutants		Total mutants		Reference
		No.	Per cent	No.	Per cent	No.	Per cent	
Untreated	769	0		1?	.13	1	.13	Horowitz *et al.* (1946)
Treated with sulfur mustard	760	17	2.2	12	1.5	29	3.8	
Untreated	1507	6	.4	1	.06	7	.46	Tatum (unpublished)
Treated with nitrogen mustard	2657	23	.86	35	.94†	48	1.80	

* Requirements of biochemical mutants:

Sulfur mustard
methionine, cystine 6
leucine 1
adenine 1
p. a. b. 2
B_1 1
amino acids 1

Nitrogen mustard
methionine, cystine 9
lysine 5
tryptophane 1 (new genetic type)
B_6 2
leucine 1
choline 1
B_1 1
uracil 2
isoleucine + valine 1 (new genetic type)
amino acids 1
yeast extract 1

† Maximum; 4.2% in one series.

in producing biochemical mutations in *Escherichia coli* (Tatum, 1947—see TABLE 2). The mutations detected in these two organisms are characterized by failure of syntheses of specific growth-factors. These failures probably result from genetically altered or inactivated enzymic capacities (Horowitz *et al.*, 1945; Beadle, 1945a).

The striking experiments of Strong and his collaborators, carried on over the past few years (Strong, 1945, 1946), have suggested that 20-methylcholanthrene, under suitable conditions (long continued treatment and inbreeding with selection), is capable of producing, in mice, somatic mutations (mosaics) and germinal mutations (susceptibility to fibrosarcoma and gastric carcinoma) inherited as dominant characters, and a variety of color mutations of both dominant and recessive types. The occurrence and behavior of many of these mutations seem to indicate a general genic instability. It is noteworthy that, in mice selected for resistance to local methylcholanthrene carcinogenesis, many types of neoplasms have been produced by the carcinogen, acting on different

TABLE 2

EFFECT OF ββ-DICHLORODIETHYLMETHYLAMINE ON *E. coli* 58

(Tatum, 1947)

Series	Treatment	Survival ratio	Cultures tested	Mutant strains obtained*							
				Stable				Unstable		Total	
				Total		Different					
				No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent
1	.05% 20 min.	$1/1\text{x}10^5$	2034	5	.25	4	.2	1	.05	6	.3
2	.1% 30 min.	$1/1.5\text{x}10^6$	1234	5	.4	4	.32	13	1.05	18	1.4
3	1.0% 30 min.	$1/2\text{x}10^7$	1182	10	.8	6	.50	1	.08	11	.93

Requirements:

Series 1

proline, 2
methionine, 1
phenylalanine, 1
tyrosine, 1

Series 2

proline, 1
phenylalanine, 1
threonine, 1
hydrolyzed casein, 2

Series 3

proline, 4
methionine, 1
pyrimidines, 1
purines, 1
adenine or hypoxanthine, 1
indole, 2

tissues. Strong has suggested that methylcholanthrene may owe its carcinogenic action to its effect in inducing somatic mutations.

Attempts are now being made in our laboratory, by Mr. R. W. Barratt and myself, to detect and analyze the mutational effects of carcinogens on *Neurospora.** The results with the more water-soluble 20-methylcholanthrene-endosuccinic acid suggest that this compound induces gene mutations in *Neurospora* similar, in effect, to those produced by irradiation and by treatment with the nitrogen and sulfur mustards, although the carcinogen is less effective than the other agents under the conditions used. Out of 3075 treated cultures, 6 biochemical mutant strains have been obtained. Dr. L. Garnjobst has found that one of these (No. Y6516) contains two apparently independent gene mutations. One gene controls the synthesis of p-aminobenzoic acid, and the other, of uracil or uridylic acid. With the very low mutation rate, the occurrence of two mutant genes in one strain is perhaps more than coincidence. In addition, neither mutation has been frequent in material treated with other agents. The p-aminobenzoicless character differs genetically from that most frequently found in irradiated material. The rather high frequency of occurrence of morphological mutant types found with methylcholanthrene in certain experiments may be significant, although further work with isogenic stocks will be needed to establish this.

A number of attempts have been made to examine the effects of certain carcinogens on growth and metabolism of microorganisms. Although the results have not been too convincing, nor very thoroughly analyzed, they should be mentioned. DeClerck (1942) has reported more or less temporary effects of the carcinogen, Styryl 430, on the respiration of yeast, either stimulatory or depressing, depending on the duration of exposure. Most recently, Reese and Reese (1945), as well as Raab (1946), have demonstrated inhibitory effects of dibenzanthracene and of phenanthrene on the growth of bacteria and certain fungi. Spencer and Melroy (1942) have reported the effect of methylcholanthrene in limiting the growth potentialities of certain bacteria and of paramecium. In none of these instances has genetic analysis been possible. About all that can be said, at present, is that these carcinogenic substances are biologically active on microorganisms.

---

* This investigation is being supported by a grant from the American Cancer Society.

It seems reasonably well established that methylcholanthrene is capable of inducing gene mutations in mice and in *Neurospora,* and that the mutations produced affect visible morphological characters as well as biochemical ones such as susceptibility to cancer, and the synthesis of enzymes concerned with the manufacture, by the cell, of nutritional factors. It seems likely that the fundamental basis of these various mutations is the same, and that germinal and somatic mutations of both types may occur in animal tissues. This would reaffirm the significance of mutation in carcinogenesis. Somatic mutations affecting enzymes and, hence, the synthesis of cellular constituents could result in the modification of the growth capacities of the cell, perhaps along the lines developed by Potter (1944) in presenting his concept of the origin of cancer as based on enzymic mutation. Somatic mutations could also result in the release of the cell from its growth limitations, perhaps by facilitating or restoring the synthesis of a limiting growth-factor (negative control by the organism—Lederberg, 1946). As examples of this might be cited the reversion of a leucine-dependent *Neurospora* mutant strain to leucine independence by back-mutation (Ryan and Lederberg, 1946) and the appearance of pleomorphic strains in cultures of *Trichophyton* as described by Robbins and Ma (1945). These pleomorphic strains differ from the original strains in being able to grow vigorously on ammonium nitrogen, and could result from mutation. Alternatively, somatic mutations might result in the production of an antagonist, or an enzyme resistant, to a hormone or other factor normally limiting enzyme activity and growth (positive control by the organism). Certain strains of *Drosophila* are characterized by decreased resistance to CO_2 (L'Héritier and Teissier, 1939). Certain color mutants of *Drosophila* also vary in their resistance to thiourea (Goldsmith and Harnly, 1946). Mutation in microorganisms leads to increased resistance to bacterial viruses (Luria, 1947; Demerec and Latarjet, 1947), to heavy metal salts (Severens and Tanner, 1945), and to penicillin (Demerec, 1945).

The effects of methylcholanthrene on tissue cultures, as reported by Earle and his collaborators (Earle, 1943), are not inherently antagonistic to a mutational effect of methylcholanthrene. The change of all cells in a culture may have resulted from strong selection for a mutant cell type, as for virus resistance in microorganisms (or, conceivably, to cell-type transformation, as in *Pneumococcus*—Avery *et al.*, 1944). The apparent continuous change in cell and tissue morphology and malignancy with longer treatment time is reminiscent of the step-wise acquisition of penicillin and virus resistance in bacteria, attributed to successive mutational steps (Demerec and Fano, 1945; Demerec, 1945). Finally, the stability of Earle's tissue culture types up to six years after cessation of exposure to methylcholanthrene, is not incompatible with a mutagenic action of the carcinogen, especially since genetic tests for gene mutation, as opposed to visible chromosomal aberrations of mitotic irregularities, are impossible in this material.

The mutational approach to carcinogenesis is also not in apparent disagreement with the known factors involved in this process. Since a mutation, germinal or somatic, gene or chromosomal, determines the potentiality of a cell and not necessarily the expression of the character, it may be suggested that environmental factors, nutrition, irritation, hormonal levels, or biochemical changes resulting from further somatic mutation or virus infection, may determine the expression of the character, in this case neoplastic growth, in cells possessing these potentialities as the result of their genetic make-up.

That exogenous carcinogens are involved, in certain instances, in experimental cancer, either directly in the induction of somatic genic or extranuclear mutations, or indirectly through modification of enzyme synthesis or potentiation of the expression of gene-determined characters, is apparently undeniable. If the occurrence of "spontaneous" neoplasms requires like stimuli, it is possible that carcinogenic substances of endogenous origin are involved (for discussion and references, see Burk and Winzler, 1944). As has been pointed out, metabolic products of sterols and bile acids, or of sex or other hormones (Dobriner *et al.*, 1946), may be involved. Irritation, physical or chemical such as CCl_4 (Edwards and Dalton, 1942) or irradiation could likewise give rise to similarly active metabolic products. Kensler, Young, and Rhoads (1942) have suggested that metabolic products of carcinogens are the ultimate active substances in carcinogenesis. Spontaneous gene mutation, either germinal, somatic, or cytoplasmic, could also lead to the increased production or to the accumulation of carcinogens (see Wachtel, 1946a & b) derived from exogenous or endogenous precursors by modifying enzymes concerned with the normal metabolism of these substances, in a manner analogous to the modification of biochemical reactions by gene mutation in *Neurospora* or inheritable metabolic disorders in man.

Nor is the virus theory of cancer opposed to this interpretation of carcinogenesis. The demonstrations of the virus nature of Shope rabbit papilloma, of Rous chicken sarcoma and, most recently, of the mouse mammary tumor milk factor (Bittner, 1942) or the virus-like principle of Taylor (1943), find their counterpart in the demonstrations of extranuclear inheritance of characters in *Pneumococcus* (Avery *et al.,* 1944), in yeast (Lindegren, 1945), in protozoa (Sonneborn, 1945), in *Drosophila* (Kalmus and Mitchison, 1946; Mampell, 1945), and in plants (Rhoades, 1943). If current views of the relation of gene to enzyme (Horowitz *et al.,* 1945; Beadle, 1945b, Spiegelman and Kamen, 1946) and of virus to gene and cell (Darlington, 1944; Woods and duBuy, 1945) are accepted, the clear-cut demarcations of gene from virus and cytoplasm from nucleus are erased. A virus could be considered as an exogenous carcinogen; also, as a mutagenic entity, if virus infection can be looked upon as a mutation.

On the other hand, a virus of endogenous origin could be thought of as arising from mutation of normal cell constituents. If we interpret mutation in the broad sense to mean a heritable change involving any level between nuclear gene and cytoplasmic effect, all of these apparently antagonistic theories of carcinogenesis fall into one general pattern.

A comparison of some of the biological and chemical effects of different carcinogenic or mutagenic influences or agents will perhaps aid in clarifying the situation in regard to mutation and cancer. If mutation is involved, mutagenic agents should be carcinogenic and carcinogenic agents should have specific and demonstrable primary effects on heredity-determining entities (cell nuclei) or on specific chemical components of these entities. These effects, to be significant, should be evident at levels of treatment below those which produce secondary effects such as cytoplasmic protein denaturation or non-specific cell disorganization. Detectable effects might be significant at different levels of cell organization (as evidenced by abnormal mitosis or chromosome division, most convincingly as shown by genic mutation or chromosomal aberration, or perhaps, basically, as indicated by modification of the physicochemical properties of nuclear material or specific constituents thereof, particularly nucleoprotein and nucleic acid).

Radiation seems to fulfill all of these requirements, since both x-ray and ultraviolet are carcinogenic and induce mutations, probably through their specific effects on nucleoprotein. Ultraviolet of 2537Å depolymerizes sodium thymonucleate very strongly (Hollaender, Greenstein, and Jenrette, 1941). X-ray has recently been shown to induce depolymerization of thymonucleohistone and sodium thymonucleate (Sparrow and Rosenfeld, 1946).

The mutagenic chemicals of the β-chloroethylamine and sulfide types have definite specific nucleotoxic actions, manifested by the effects of low concentrations in inhibiting mitosis and in inducing mutations of different types in *Tradescantia, Drosophila,* bacteria, and *Neurospora.* No specific action of these compounds on nucleoprotein or thymonucleate has been reported yet, nor have rigorous tests for carcinogenicity been described.

Carcinogenic chemicals such as methylcholanthrene and dibenzanthracene seem specifically to affect nuclear and chromosome behavior, as evidenced by their effects on mitosis, as well as on nuclear and chromosome size (Biesele, 1944), similar to the effect of camphor on yeast. The effect of the carcinogens or their metabolic products may be on nucleoproteins, in view of the adsorption of such substances as benzpyrene by mitochondria of *Daphnia* and mammalian cells (Graffi, 1940). This observation is of particular interest in relation to the theory of the origin of viruses from mitochondria proposed by Graffi and by Woods and duBuy (1945), and in relation to the role of mitochondrial mutation in carcinogenesis suggested by these authors. The results with mice and with *Neurospora* suggest that at least one carcinogen, methylcholanthrene, has mutational effects.

Another action of all these agents is that of inhibiting enzyme production or activity. The latter effect of carcinogens, as shown by *in vitro* experiments (Kensler *et al.,* 1942; Potter and du Bois, 1943, is not necessarily directly related to carcinogenicity. The demonstrations of Greenstein and Chalkley (1945) that thymonucleate combines with and regulates the activity of respiratory enzymes, suggests that the effect of carcinogens on enzyme activity *in vivo* (see Potter, 1944) might be indirect, and that the primary effect might be on nucleoproteins, nuclear or extranuclear. In either case, the effect might be heritable and, in the broad sense, constitutes a mutation. The inhibitions of enzyme activity by irradiation and by the mustards (see Herriott *et al.,* 1946) may be evident only at higher intensity or concentration levels than those giving mutational effects.

In conclusion, a critical evaluation of the evidence for and against the mutational theory of cancer, perhaps, cannot yet be made. Nevertheless, many of the known facts regarding the changes taking place during carcinogenesis seem to support this theory, and perhaps none are in actual disagreement. As indicated previously, if a broad definition of mutation is accepted, the theory that cancer results from mutation may, in time, dovetail with other theories, viral, adaptational, and enzymatic, that have been proposed, and may prove to be the common factor in carcinogenesis.

Note added in proof. Dr. M. Demerec has recently reported that four carcinogens, 1,2,5,6 dibenzanthracene, methylcholanthrene, benzpyrene, and β-naphthylamine, induce mutations in *Drosophila* (Nature **159**:604, 1947; and a paper presented at the 1947 meeting of the National Academy of Sciences, abstract in Science **105**:634, 1947).

BIBLIOGRAPHY

Auerbach, C., & J. M. Robson
1946. Nature **157**: 302.

Avery, O. T., C. M. MacCleod, & M. McCarty
1944. J. Exp. Med. **79**: 137.

Bauch, R.
1941. Naturwiss. **29**: 687.

Beadle, G. W.
1945a. Physiol. Rev. **25**: 643
1945b. Chem. Rev. **37**: 15.

Biesele, J. J.
1944. J. Nat. Cancer Inst. **4**: 373.

Bittner, J. J.
1942. Science **95**: 462.

Burk, D., & R. J. Winzler
1944. Ann. Rev. Biochem. **13**: 487.

Darlington, C. D.
1944. Nature **154**: 164.

de Clerck, J.
1942. Bières et Boissons **3**: 176.

Demerec, M.
1945. Ann. Mo. Bot. Garden **32**: 131.

Demerec, M., & V. Fano
1945. Genetics **30**: 119.

Demerec, M., & R. Latarjet
1946. Cold Spring Harbor Symposia Quant. Biol. **11**: 38.

Dixon, M., & D. M. Needham
1946. Nature **158**: 432.

Dobriner, K., C. P. Rhoads, S. Lieberman, B. R. Hill, & L. F. Fieser
1946. Science **99**: 494.

Earle, W. R.
1943. J. Nat. Cancer Inst. **4**: 165, 213.

Edwards, J. E., & A. J. Dalton
1942. J. Nat. Cancer Inst. **3**: 19.

Furth, J., M. C. Boon, & N. Kaliss
1944. Cancer Res. **4**: 1.

Gilman, A., & F. S. Philips
1946. Science **103**: 409.

Goldsmith, E. D., & M. H. Harnly
1946. Science **103**: 649.

Graffi, A.
1940. Z. Krebsf. **50**: 196, 501.

Greenstein, J. P.
1945. Ann. Rev. Biochem. **14**: 643.

Greenstein, J. P., & H. W. Chalkley
1945. Ann. Mo. Bot. Garden **32**: 179.

Hadorn, E., & H. Niggli
1946. Nature **157**: 162.

Herriott, R. M., M. L. Anson, & J. H. Northrop
1946. J. Gen. Physiol. **30**: 185.

Hollaender, A., J. P. Greenstein, & W. V. Jenrette
1941. J. Nat. Cancer Inst. **2**: 23.

Horowitz, N. H., D. Bonner, H. K. Mitchell, E. L. Tatum, & G. W. Beadle
1945. Am. Naturalist **79**: 304.

Horowitz, N. H., M. B. Houlahan, M. G. Hungate, & B. Wright
1946. Science **104**: 233.

Kalmus, H., & A. N. Mitchison
1946. Nature **157**: 230.

Kensler, C. J., N. F. Young, & C. P. Rhoads
1942. J. Biol. Chem. **143**: 465.

Law, L. W.
1938. Proc. Nat. Acad. Sci. **24**: 546.

Lederberg, J.
1946. Science **104**: 428.

Levan, A.
1946. Nature **158**: 626.

Levan, A., & C. G. Sandwell
1943. Hereditas **29**: 164.

L'Héritier, P., & G. Teissier
1939. Proc. 7th Internat. Genetics Congress, Abstract **172**: 190.

Lindegren, C. C.
1945. Ann. Mo. Bot. Garden **32**: 107.

Luria, S. E.
1946. Cold Spring Harbor Symposia Quant. Biol. **11**: 130.

Mampell, K.
1945. Genetics **30**: 496.

Potter, V. R.
1944. Adv. Enzymology **4**: 201.

Potter, V. R., & K. P. Du Bois
1943. J. Gen. Physiol. **26**: 391.

Raab, W.
1946. Science **103**: 670.

Reese, E., & L. Reese
1945. Growth **9**: 177.

Rhoades, M. M.
1943. Proc. Nat. Acad. Sci. **29**: 327.

Robbins, W. J., & R. Ma
1945. Am. J. Bot. **32**: 509.

Ryan, F. J., & J. Lederberg
1946. Proc. Nat. Acad. Sci. **32**: 163.

Severens, J. M., & F. W. Tanner
1945. J. Bact. **49**: 383.

Sonneborn, T. M.
1945. Ann. Mo. Bot. Garden **32**: 107.

Sparrow, A. H., & F. M. Rosenfeld
1946. Science **104**: 245.

Spencer, R. R., & M. B. Melroy
1942. J. Nat. Cancer Inst. **3**: 1.

Spiegelman, S., & M. D. Kamen
1946. Science **104**: 581.

Steinberg, R. A., & C. Thom
1940. J. Heredity **31**: 61.

Stier, T. J. B., & J. G. B. Castor
1941. J. Gen. Physiol. **25**: 229.

Strong, L. C.
1945. Proc. Nat. Acad. Sci. **31**: 290.

Strong, L. C.
1946. Yale J. Biol. Med. **18**: 359.

Tatum, E. L.
1946. Cold Spring Harbor Symposia Quant. Biol. **11**: 278.

Taylor, A.
1943. Science **97**: 123.

Thomas, J. A., & S. Chevais
1943. C. R. Soc. Biol. **137**: 187.

Wachtel, H. K.
1946a. Science **103**: 556.
1946b. Nature **158**: 98.

Woods, M. W., & H. G. Du Buy
1945. Science **104**: 469.

THE NUTRITION OF MONOCELLULAR ANIMAL ORGANISMS*

By GEORGE W. KIDDER

Amherst College, Amherst, Massachusetts

The justification for the material I am about to present was implied, I believe, at the opening of the conference. Any advances we can make in our understanding of fundamental problems of cell metabolism will surely be of some aid to a better understanding of the cancer problem.

Perhaps the ideal of the animal nutritionist is, first, a complete biochemical analysis of the nutritional requirements of the organisms used for experimentation and, second, an analysis of the physiological role of the various components of the diet. Although our knowledge of the function of many of the components of animal diets has increased enormously within the past few years, there are still more gaps to fill than have been filled up to the present. Too often, this type of work is seriously handicapped by a biotic environment too complex to analyze. Just as often, in the past, this complexity has not even been appreciated. The nutritional-requirement problem suffers most in this respect, for it becomes very difficult indeed to decide whether the animal in question does or does not require a certain substance, when the intestinal tract of the animal is harboring unknown microorganisms with varying synthetic abilities. The ideal situation with which to work, therefore, would be aseptic laboratory animals. Aseptic vertebrates have been established and maintained, but the addition to our knowledge of nutrition by the use of these animals has been almost nil. It appears that the persistence of the animal nutritionists in using rats, mice, chicks, and dogs has been maintained because these organisms are more nearly like man and his domestic animals and, therefore, what knowledge can be had from such studies will have a better chance of being applied directly. This tendency has been prevalent also in fields other than nutrition.

If we set as our goal the establishment of pure cultures, *i.e.*, experimental animals with which we can be sure that our analyses pertain to the single species, then it seems that a number of criteria should be considered. First, will the animal grow and carry on its metabolic activities in a relatively normal manner when freed of its associates? Second, will its size be such that it can be readily manipulated under aseptic conditions? Third, will the bulk of its nutrients not be hopelessly expensive?

---

* A portion of the results reported here were obtained with support from a grant from The U. S. Public Health Service. I should like to acknowledge the valuable help I have received, in much of this work, from a number of my students, and especially from Dr. Virginia C. Dewey, who is again with me as Research Assistant at Amherst College.

Fourth, it would be desirable to be able to compare the results obtained with other types of animals for practical as well as theoretical reasons. This last consideration comes down to the question: How different are animals fundamentally? Are there basic nutritional patterns which are really animal in character? If so, there is no reason why not to examine the Protozoa, since some of them seem to meet the criteria mentioned. They certainly have not proved to be simple, as a group. Most of those studied so far have defied complete analysis. It has been possible to show that a number of the vitamins important to higher organisms are also needed by many Protozoa, but the requirements for specific amino acids have been analyzed for one genus only. So far, the majority of the pure-culture protozoans have been grown in mixtures containing natural products such as proteins, peptones, various extracts of plant and animal origin, etc., and the situation here is only a little better than in non-sterile cultures.

Time does not permit the development of the subject of protozoan nutrition through the stages of sheer speculation, when other microorganisms in the medium were thought to be unimportant, through the not too distant past when nutritional studies were based upon the types and brands of peptones used, to the present when more precise methods are utilized. Credit for our present appreciation of many of the problems in this field goes to André Lwoff of the Pasteur Institute, who, in 1923,[1] announced the establishment of bacteria-free cultures of a protozoan of typical animal nature. The various strains of this genus of ciliate, *Tetrahymena,* have proved invaluable for nutritional studies.

Although *Tetrahymena* is the only true protozoan which has been cultured in the absence of natural proteins or peptones, we have some information regarding the growth factor requirement of other types. The flagellates listed in TABLE 1 have been cultured only in the presence of peptones with or without serum, so that the column, "unidentified factors", should be filled in in every case. In this table, however, the plus sign indicates special factors, perhaps peculiar to the species studied. The blanks simply indicate that nothing is known regarding that factor, while the minus sign indicates that an exogenous source of the factor is unnecessary. The trypanosomids of insects apparently synthesize adequate amounts of cholesterol and ascorbic acid. Half of those studied require hematin, while the necessity of an exogenous source of thiamine has been found in all those tested for this vitamin. The trichomonads, leishmanias, and *Schizotrypanum* all require serum factors, so that determinations of a possible thiamine requirement have not been made. All of the trichomonads seem to possess the same growth factor pattern, as far as tested. The leishmanias from reptiles differ from the human parasites, and from *Schizotrypanum,* in their lack of an ascorbic acid requirement.

TABLE 1

GROWTH FACTOR REQUIREMENT: ANIMAL FLAGELLATE

Organism	Growth factors					Reference
	Cholesterol	Ascorbic acid	Hematin	Thiamine	Unidentified factors	
Strigomonas oncopelti	—	—	—	+		M. Lwoff (2–7)
Strigomonas ctenocephali	—	—	+			
Strigomonas fasciculata	—	—	+	+		
Strigomonas culicidarum anophelis	—	—	+	+		
Strigomonas culicidarum culcis	—	—	—	+		
Strigomonas muscidarum	—	—	+			
Strigomonas parva	—	—	—			
Strigomonas media	—	—	—			
Trichomonas batrachorum	+		—		+	Cailleau (8–17)
Trichomonas columbae	+	+	—		+	
Trichomonas foetus	+	+	—		+	
Eutrichomastix colubrorum	+	+	—		+	
Leishmania agamae		—	+		+	M. Lwoff (18, 19)
Leishmania ceramodactyli		—	+		+	
Leishmania tropica		+	+		+	
Leishmania donovani		+	+		+	
Schizotrypanum crusi		+	+		+	

One member of the Sarcodina has been studied in pure culture (TABLE 2). *Acanthamoeba* requires thiamine or the thiazole and pyrimidine parts of the molecule. Later work by Lwoff[27] suggests that the pyrimidine portion alone is sufficient.

The recent work by Ball and his collaborators[21] tends to show that *Plasmodium* requires rather high concentrations of p-aminobenzoic acid, in addition to a number of unidentified factors.

Of the Ciliates other than *Tetrahymena*, very little of an exact nature is known. *Glaucoma scintillans* requires thiamine,[22] in addition to unidentified factors present in yeast and peptones. None of the known vitamins were determined for *Colpidium*, due to the nature of the unknown factors required.[23] The work of Johnson and Tatum[24] shows that *Paramecium* requires at least two unidentified factors from plasmolyzed bacteria, one extremely labile and the other stable to heat. Working with this "plasmoptysate," Tatum and collaborators[25, 26] have shown that *Colpoda* requires, in addition, rather high concentrations of thiamine, pantothen, riboflavin, nicatinamide, and pyridoxine, but does not require an exogenous source of PAB, biotin, or inositol.

In no true Protozoa are the amino acid requirements known, except for the genus *Tetrahymena*, and it is this genus that I want to describe

TABLE 2

GROWTH FACTOR REQUIREMENT

Sarcodinia

Organism	Growth factors									Reference
	Thiamine	Pantothen	Riboflavin	Nicotinamide	Pyrodoxine	p-Aminobenzoic acid	Biotin	Inositol	Unidentified factors	
Acanthamoeba castellanii	+									A. Lwoff (20)
					Sporozoa					
Plasmodium knowlesi						+			+	Ball, *et al.* (21)
					Ciliates					
Glaucoma scintillans	+								+	Kidder & Dewey (22)
Colpidium campylum									+	Peterson (23)
Paramecium multimicro-nucleata									+	Johnson & Tatum (24)
Colpoda duodenaria	+	+	+	+	+	—	—	—	+	Tatum, *et al.* (25, 26)

in some detail. Indeed, it may be said that this is the only genus in the animal kingdom, *grown under aseptic conditions,* in which the amino acid requirements have been reported in the literature. A possible exception to this statement is the avian embryonic tissues reported this year by White, but no analysis of the actual requirements for individual amino acids was given. Very recently, Dr. Schultz has succeeded in this respect with *Drosophila* (personal communication).

For some years now, we have been studying the biochemistry of six strains of *Tetrahymena geleii* and two strains of *Tetrahymena vorax.* The general nutritional patterns are the same for all strains, although differing in details. What I have to say from now on applies specifically to *Tetrahymena geleii,* strain W.

I shall review, briefly, the growth factor or vitamin pattern for *Tetrahymena,* which has been established except for one unknown factor (Factor II). The only member of the vitamin B complex which this ciliate *must* have from an outside source is pteroylglutamic acid (TABLE 3).

TABLE 3

GROWTH FACTOR PATTERN: *Tetrahymena geleii W*

	Growth rate increased	Maximum yield increased
I. Required		
A. Pteroylglutamic acid or Vitamin B conjugate		
B. Purine (especially guanine)		
C. Pyrimidine (especially cytidylic acid)		
D. Factor II (unidentified)		
II. Synthesized in inadequate amounts for optimum growth*		
A. Thiamine	0	+
B. Riboflavin	+	+
C. Pantothen	0	+
D. Niacin	+	+
E. Pyrimidine	+	+
F. Biotin	+	+
III. Synthesized in adequate amounts†		
A. p-Aminobenzoic acid		
B. Inositol		

* Actual vitamin synthesis was shown by microbiological methods using *Lactobacillus casei*.
† Both p-aminobenzoic acid and inositol were shown to be synthesized by the ciliates, when employing assay methods with "aminobenzoic-less" and "inositol-less" mutants of *Neurospora crassa*. These strains of *Neurospora* were furnished by Dr. E. L. Tatum of Yale University.

Quantitative determinations show that its requirement is approximately ten times greater than that of the PGA-requiring bacteria.[28] Recently, we received a sample of vitamin B_c conjugate from Dr. Bird of the Parke-Davis Laboratories, and we find that *Tetrahymena* possesses the necessary conjugase for this substance. The polyglutamate was twice as active as the monoglutamate. The *Streptococcus lactis* R factor, obtained from the Merck Laboratories, could partially substitute for PGA at high levels,[28] but pteroic acid, supplied by Dr. Stokstad of the Lederle Laboratories, was entirely without activity. This may be direct evidence that the SLR factor and pteroic acid are not identical, although they appear to be when tested on bacteria. It is more probable that the SLR factor used was contaminated with PGA, since it was a natural product.

Purines are required for growth, but the ciliate is not limited to the use of one. The most active is guanine (or its nucleoside or nucleotide), but adenylic acid alone is sufficient for indefinitely transplantable growth.[29]

Either uracil (or its nucleoside or nucleotide) or cytidylic acid can be used to supply the pyrimidine requirement, the latter being the most active. Hydrolyzed nucleic acid can, of course, meet the needs for both the purines and pyrimidines, and this is regularly used for many of the studies.[29]

What we have called Factor II[29, 30, 31] represents the only unidentified substance required by *Tetrahymena*. This factor is under investigation, at present, in Dr. Stokstad's laboratory, but it should be stated that its chemical characteristics are such that it has not hindered the work on the other growth factors, from all of which it is separable.[31, 32]

Six members of the B-complex can be synthesized by the ciliate, but at rates too low for optimum growth. The addition of thiamine does not increase the growth rate during the logarithmic phase, but raises the maximum yield and longevity of the culture.[33] Another interesting point regarding thiamine is that individual cell size is reduced to as much as one quarter of normal when the vitamin is produced endogenously.

The other vitamins in this category, with the exception of pantothen, not only increase the maximum yield when added to the medium, but the growth rate as well. Exogenous pantothen does not affect the growth rate.[32]

p-Aminobenzoic acid and inositol are synthesized in appreciable quantities by the ciliate, and added amounts are without effect. The exact status of choline is still in doubt.

In all studies on nitrogen and carbon metabolism, the substances listed in groups I and II are included, together with choline and the inorganic salt mixtures used by Hall and Cosgrove.[34]

The usual procedure was followed in determining the essential amino acids for the ciliate, once it was possible to obtain growth in protein hydrolysates. A mixture of the 19 amino acids found in casein was used and single omissions were tested.[35] From these results, as shown in TABLE 4, it appears that the ten amino acids regarded as being indispensable for mammals were either not synthesized or synthesized in amounts inadequate for optimum growth. Glycine appeared to fall into the same category. Further tests, however, soon proved that the results

TABLE 4

Tetrahymena geleii W

Amino acid omitted	Cells/ml., 3rd transplant	Amino acid omitted	Cells/ml., 3rd transplant
None	320,000	Leucine	0
Alanine	270,000	Lysine	0
Arginine	6,500	Methionine	0
Aspartic acid	290,000	Phenylalanine	0
Cysteine	240,000	Proline	210,000
Glutamic acid	200,000	Serine	300,000
Glycine	260	Threonine	0
Histidine	0	Tryptophane	0
Hydroxyproline	240,000	Tyrosine	210,000
Isoleucine	0	Valine	0

were misleading regarding glycine. This seems to function as a detoxifying agent for some of the non-essentials included,[30] for optimum growth was possible using just the other 10, with small amounts of serine added. Serine was the only one of the non-essential amino acids which was really stimulatory, and has been incorporated into the medium for the later work. It seems possible that the role of glycine in chick nutrition, where it is reported to be essential in nature, is similar to what we have found here.

Our eventual amino acid mixture, therefore, consisted of the ten essential amino acids with serine added, the proportions being based on the published analysis of gelatin, with arbitrary amounts of tryptophane and valine (TABLE 5). The final medium contained 592.11 micrograms of nitrogen per ml. Using this medium and varying the individual amino acids, we determined the optimum concentrations for each. These figures (shown in the second column of TABLE 5) which were finally used for the *proportions,* were considerably lower than in the original mixture. Then, of course, the total nitrogen (71.45 micrograms/ml.) was too low, but by increasing the total concentration eight times, and maintaining the same proportions, optimum growth resulted.

TABLE 5

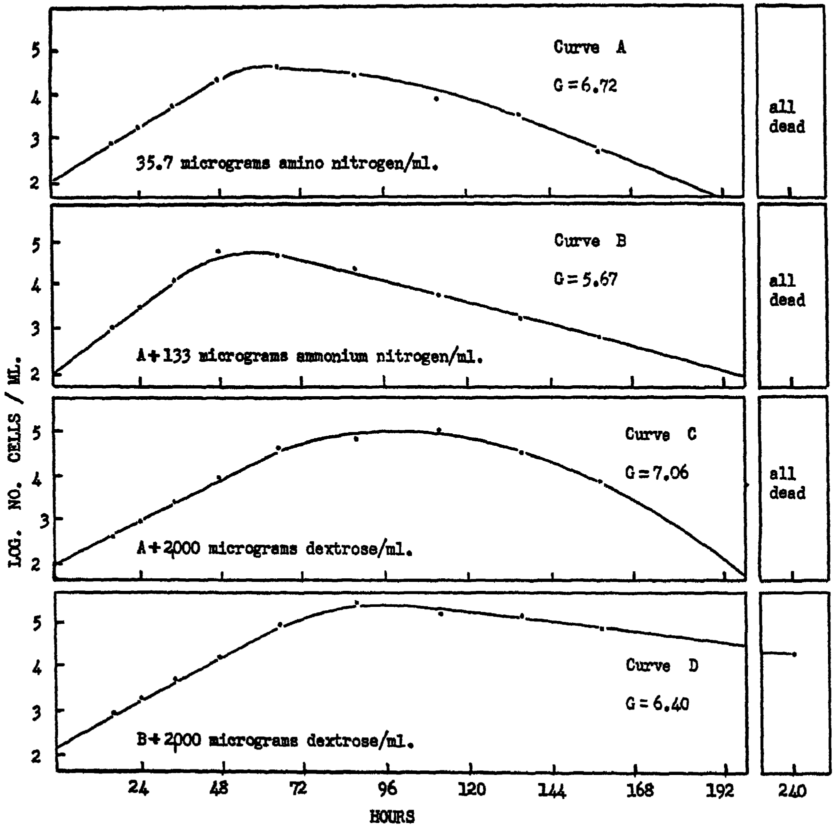
	Micrograms/ml.	
	Based on gelatin analysis	Based on growth optimum
1(+)-arginine mono-hydrochloride	820	80
1(—)-histidine mono-hydrochloride	100	25
dl-isoleucine	350	10
1(—)-leucine	350	50
dl-lysine	600	60
dl-methionine	340	50
dl-phenylalanine	140	35
dl-threonine	200	25
1(—)-tryptophane	100*	5
dl-valine	200*	15
dl-serine	40	60
Total	3240 (N 592.11)	415 (N 71.45)

* Amount arbitrarily added.

During the course of this work, we compared the activity of natural and racemic tryptophane and leucine. The dl-tryptophane was found to be approximately 75 per cent as active as the natural material. It seems that the unnatural isomer can be metabolized, but at a low rate. On the other hand, natural leucine was approximately six times as active as the racemic mixture. This means that the unnatural isomer will not substitute for the natural and is, in addition, inhibitory.

A number of substances were tested for their sparing action for specific amino acids. Cysteine was found to spare, but not replace, methionine. Homocystine can replace methionine (choline was present). Tyrosine can spare, but not replace, phenylalanine. The ciliate lacks the ability to condense indole with serine for the synthesis of tryptophane. Serine, cystine, and glycine together can partially replace valine.

A number of years ago, Schoenheimer and his collaborators[37, 38] were able to show, by the use of the tracer technique, that ammonium nitrogen was used by the rat in the synthesis of some of its amino acids. Rats were fed ammonium citrate containing N^{15}, and later analysis revealed the presence of the isotope in some of their amino acids. It was of



G = generation time in hours during the logarithmic growth phase

FIGURE 1.

interest, therefore, to determine whether or not the ciliate could utilize ammonium nitrogen and, if so, whether or not it could replace, to any appreciable extent, the amino nitrogen needed for protoplasmic synthesis. Accordingly, the concentration of the eleven amino acids employed was reduced to $\frac{1}{16}$ the optimum level. To this, varying concentrations of ammonium pyruvate were added, with and without dextrose. A typical example of the results of one set of concentrations is shown in FIGURE 1. When the organisms are limited to this small amount of amino acids, for architectural material and the synthesis of the other amino acids which go to make up their protein, as well as their sole energy supply, the life of the culture is curtailed (curve A). During the death phase, the cells become extremely emaciated. When ammonium pyruvate is added in the concentration shown in B (1,000 micrograms/ml. containing 133 micrograms of N/ml.), the caloric content of the medium is increased approximately five-fold by the pyruvate, but pyruvic acid is utilized so readily that the only striking difference is the increase in growth rate and maximum yield. Death results at about the same period. When 2,000 micrograms/ml. of dextrose is added to the amino acid medium (C), the caloric content is increased ten-fold. Growth is slower, but the maximum yield is increased. Probably, essentially all of the amino acids can be used for protein synthesis, either as the essential building blocks or for synthesis of the non-essentials, and none need be consumed for energy. By ten days, however, these cultures are likewise dead (curve C). Evidence of the wasting away of the cells is postponed. When both ammonium pyruvate and dextrose are added (D), the caloric content of the medium is approximately fifteen times greater than that supplied by the amino acids alone. The shape of the growth curve is normal, the cells do not become shriveled, and a fairly high population is maintained for long periods of time. The addition of ammonium nitrogen makes it possible for the ciliates to synthesize protoplasm in a medium otherwise extremely deficient. The conditions, then, for the maximum utilization of ammonium nitrogen by the ciliates, are an adequate energy source, just enough of the essential amino acids (and in the proper balance) to meet the immediate needs of body protein synthesis, and an adequate supply of ammonium nitrogen for the synthesis of the non-essential amino acids.

We feel that these results may have far-reaching nutritional implications. It was shown as early as 1900, by Kellner,[39] that ruminants could utilize ammonium nitrogen as part of their diet, but he ascribed this to the ability of the microorganism in the rumen to convert the ammonia to protein. This idea was given support, in 1939, by Schmid,[40] who found that, of the microorganisms present in the rumen, only certain of the bacteria were responsible for converting ammonium nitrogen to bacterial proteins. Essentially the same results were reported by Klein, in 1941.[41] The possibility, indicated clearly by Schoenheimer's work with tracer nitrogen in the rat, that higher animals, other than the herbivores, might be able to utilize non-amino-nitrogen in a nutritional sense, does not seem to have received the attention it deserves. If it is at all legitimate to compare the metabolism of *Tetrahymena* with higher types on the basis of its amino acid requirements, its vitamin pattern, and, as Thomas[42] has shown, of its anaerobic respiratory mechanism, then it seems that the possibility should be explored as to whether or not an appreciable part of the nitrogen requirement of even the human can be supplied by ammonium salts—in view of their cheapness.

This brief résumé of some of the biochemical and physiological aspects of the animal protozoa, mainly *Tetrahymena,* indicates some of the basic principles which can be attacked using this type of material. If even a few of the fundamentals of animal cell metabolism can be finally understood by their use, these protozoa will have contributed their fair share.

BIBLIOGRAPHY

1. **Lwoff, André**
   1923. Sur la nutrition des infuriores. C. R. Acad. Sci. **176:** 928.
2. **Lwoff, M.**
   1932. Nature de la substance du sang indispensable à la nutrition de *Leptomonas ctenocephali* Fanth. C. R. Soc. Biol. **110:** 891.
3. **Lwoff, M.**
   1933. Recherches sur la nutrition des trypanosomides. Ann. Inst. Pasteur **51:** 55.
4. **Lwoff, M.**
   1933. Remarques sur la nutrition des trypanosomides et des bactéries parahémotrophes. Le "fer actif" de Baudisch. Ann. Inst. Pasteur **51:** 707.
5. **Lwoff, M.**
   1935. Le pouvoir de synthèse des trypanosomides des culcides. C. R. Soc. Biol. **119:** 969.
6. **Lwoff, M.**
   1937. L'aneurine, facteur de croissance pour le flagellé trypanosomide *Strigomonas oncopelti.* C. R. Soc. Biol. **126:** 771.
7. **Lwoff, M.**
   1938. L'aneurine, facteur de croissance pour les *Strigomonas* (flagellés trypanosomides). C. R. Soc. Biol. **128:** 241.
8. **Cailleau, R.**
   1936. Le cholesterol, facteur de croissance pour le flagellé *Trichomonas columbae.* C. R. Soc. Biol. **121:** 424.
9. **Cailleau, R.**
   1936. L'activité de quelques sterols envisagés comme facteurs de croissance pour le flagellé *Trichomonas columbae.* C. R. Soc. Biol. **122:** 1027.
10. **Cailleau, R.**
    1937. La nutrition des flagellés tetramitides. Les sterols, facteurs de croissance pour les trichomonades. Part 1. Ann. Inst. Pasteur **59:** 137.
11. **Cailleau, R.**
    1937. La nutrition des flagellés tetramitides. Les sterols, facteurs de croissance pour les trichomonades. Part 2. Ann. Inst. Pasteur **59:** 293.
12. **Cailleau, R.**
    1938. Le cholesterol et l'acide ascorbique, facteurs de croissance pour le flagellé tetramitide *Trichomonas foetus* Riedmuller. C. R. Soc. Biol. **127:** 861.

13. **Cailleau, R.**
1938. L'acide ascorbique et le cholesterol, facteurs de croissance pour le flagellé *Eutrichomastix colubrorum.* C. R. Soc. Biol. **127**: 1421.

14. **Cailleau, R.**
1939. L'acide ascorbique, facteur de croissance pour le flagellé *Trichomonas columbae.* C. R. Soc. Biol. **130**: 319.

15. **Cailleau, R.**
1939. Le cholesterol, facteur de croissance pour le flagellé *Trichomonas batrachorum.* C. R. Soc. Biol. **130**: 1089.

16. **Cailleau, R.**
1939. L'activité de quelques substances voisines de la vitamine C, envisagés comme facteurs de croissance pour le flagellé *Eutrichomastix colubrorum.* C. R. Soc. Biol. **131**: 964.

17. **Cailleau, R.**
1940. La nutrition de *Trichomonas gallinarum;* l'acide ascorbique, facteur de croissance. C. R. Soc. Biol. **134**: 32.

18. **Lwoff, M.**
1938. L'aneurine et l'acide ascorbique, facteurs de croissance pour le flagellé *Schizotrypanum cruzi.* C. R. Acad. Sci. **206**: 540.

19. **Lwoff, M.**
1939. Le pouvoir de synthèse des Leishmanies. C. R. Soc. Biol. **130**: 406.

20. **Lwoff, André**
1938. La synthèse de l'aneurine par le protozoaire, *Acanthamoeba castellani.* C. R. Soc. Biol. **128**: 455.

21. **Ball, E. G., C. B. Anfinsen, Q. M. Geiman, R. W. McKee, & R. A. Ormsbee**
1945. *In vitro* growth and multiplication of the malaria parasite, *Plasmodium knowlesi.* Science **101**: 542.

22. **Kidder, G. W., & V. C. Dewey**
1942. The biosynthesis of thiamine by normally athiaminogenic microorganisms. Growth **6**: 405.

23. **Peterson, R. E.**
1942. Essential factors for the growth of the ciliate protozoan, *Colpidium campylum.* J. Biol. Chem. **146**: 537.

24. **Johnson, W. H., & E. L. Tatum**
1945. The heat-labile growth factor for *Paramecium* in pressed yeast juice. Arch. Biochem. **8**: 163.

25. **Tatum, E. L., L. Garnjobst, & C. V. Taylor**
1942. Vitamin requirements of *Colpoda duodenaria.* J. Cell. & Comp. Physiol. **20**: 211.

26. **Garnjobst, L., E. L. Tatum, & C. V. Taylor**
1943. Further studies on the nutritional requirements of *Colpoda duodenaria.* J. Cell. & Comp. Physiol. **21**: 199.

27. **Lwoff, André**
1938. Les facteurs de croissance pour les microorganismes. Ann. Inst. Pasteur **61**: 580.

28. **Kidder, G. W., & R. C. Fuller**
1946. The growth response of *Tetrahymena geleii* W to folic acid and to the *Streptococcus lactis* R factor. Science **104**: 160.

29. **Kidder, G. W., & V. C. Dewey**
1945. Studies on the biochemistry of *Tetrahymena.* V. The chemical nature of Factors I and III. Arch. Biochem. **8**: 293.

30. **Dewey, V. C.**
1944. Biochemical factors in the maximal growth of *Tetrahymena.* Biol. Bull. **87**: 107.

31. **Kidder, G. W., & V. C. Dewey**
1945. Studies on the biochemistry of *Tetrahymena*. IV. Amino acids and their relation to the biosynthesis of thiamine. Biol. Bull. **89:** 131.

32. **Kidder, G. W., & V. C. Dewey**
1945. Studies on the biochemistry of *Tetrahymena*. VI. Riboflavin, pantothen, biotin, niacin, and pyridoxine in the growth of *T. geleii* W. Biol. Bull. **89:** 229.

33. **Kidder, G. W., & V. C. Dewey**
1944. Thiamine and *Tetrahymena*. Biol. Bull. **87:** 121.

34. **Hall, R. P., & W. B. Cosgrove**
1944. The question of the synthesis of thiamin by the ciliate, *Glaucoma piriformis*. Biol. Bull. **86:** 31.

35. **Kidder, G. W., & V. C. Dewey**
1945. Studies on the biochemistry of *Tetrahymena*. I. Amino acid requirements. Arch. Biochem. **6:** 425.

36. **Kidder, G. W., & V. C. Dewey**
1945. Studies on the biochemistry of *Tetrahymena*. III. Strain differences. Physiol. Zool. **18:** 136.

37. **Rittenberg, D., R. Schoenheimer, & A. S. Keston**
1939. Studies in protein metabolism. IX. The utilization of ammonia by normal rats on a stock diet. J. Biol. Chem. **128:** 603.

38. **Schoenheimer, R., S. Ratner, & D. Rittenberg**
1939. The metabolic activity of body proteins investigated with l(—)-leucine containing two isotopes. J. Biol. Chem. **130:** 703.

39. **Kellner, O.**
1900. Untersuchungen ueber den Einfluss des Asparagins und Ammoniaks auf den Eiweissumsatz der Wiederkaeuer. Z. Biol. **21:** 313.

40. **Schmid, H.**
1939. Mikrobiologische Beobachtungen an Verdauungsvorgaengen im Wiederkaeuer, insbesondere bei der Umwandlung stickstoffhaltiger Verbindungen (unter Verwendung der Ninhydrinreaktion). Z. Tierzucht. Zuechtungsbiol. **43:** 239.

41. **Klein, W.**
1941. Futterungsversuche an Milchkuhen und einem Jungrind mit Ammoniumverbindungen bei der "zymogenen Symbiose". Z. Tierzucht. Zuechtungsbiol. **48:** 277.

42. **Thomas, J. O.**
1942. The Anaerobic Carbohydrate Metabolism of *Tetrahymena geleii*. Dissertation—Stanford University.

THE NUTRITION OF MALIGNANT TISSUE *IN VITRO*

By PHILIP R. WHITE

Institute for Cancer Research, Philadelphia, Pennsylvania

The title of this paper is not of my own choosing but was suggested to me by the conference committee. I shall, I fear, fulfil my obligation more in the breach than in the fact, for there is very little of *fact* that can be said about the nutrition of any tissue, any animal tissue at least, *in vitro,* and still less about malignant tissue. All that I can do is to outline some of the problems and suggest approaches likely to lead to their solution, from my work on tissue cultures and their nutrition, first with plant tissues (*vide* Dr. Robbins) and, more lately, with animal tissues. The history of tissue cultures is a long and complex one. Haberlandt first outlined the problem in 1902, and made the first consistent attempts to solve it. The first successful short-term cultures were carried out five years later, by Harrison, using clotted lymph as both substratum and nutrient. Burrows substituted plasma for lymph, and Carrel introduced the use of embryo juice, thus making possible the consummation of cultures of unlimited duration. The Carrel method remained standard from 1912 until his death.

Obviously, a nutrient made up of blood plasma and embryo juice, even though one may be able to prepare it with sufficient precision to obtain duplicable results, is nevertheless not suitable for nutritional studies. If we are interested in thiamin nutrition, for example, we cannot start from scratch. Embryo juice contains thiamin. If we subject it to sufficiently drastic treatment to destroy or remove the thiamin, we have reason to believe that we may, at the same time, destroy or remove other nutrient factors as well. If we do not so treat it, we are dealing with an unknown base line. It is like trying to study the physics of bounce in a rubber ball without knowing if it is being bounced against a concrete floor or a deep pile carpet. If we are interested, as one of my friends has been, in the synthesis of thyroxin in tissue culture, we encounter a similar problem, for thyroxin contains iodine, while embryo juice and plasma contain minimal, only imprecisely known, and probably variable amounts of iodine. If thyroid tissue cultures fail to form thyroxin, we do not know if this is due to some general change in metabolism or merely to iodine deficiency. Mere adding of iodine to the plasma still gives us an unknown base. It is quite impossible, at present, to study the nutrition of normal tissues *in vitro* in any precise manner. Obviously, if we cannot study the nutrition of normal tissues *in vitro,* we have little chance of studying that of malignant ones. It will, thus, be necessary, before any real studies can be made on the nutrition of either healthy or malignant tissues, to develop nutrients of precisely known constitution which contain all the elements necessary for survival and continued growth and which do not contain elements which are not necessary for survival and growth. Only when it is possible to bring about failure of a culture by omission of any single element or known substance from the nutrient, and bring about marked change in behavior by addition of any single significant element or substance, can we begin to study tissue nutrition. Such a nutrient should supply the base line, the control for all studies. At the moment, however, it is not available.

A problem of this sort can be attacked in either of two ways, or both ways at once. We can analyze a complex nutrient such as embryo juice, which is known to be capable of supporting unlimited growth, to determine which of its constituents are really essential. Or we can use the scraps of evidence available from collateral fields as a basis on which to bring together a variety of relatively simple substances, in the hope of arriving at some combination which will support growth. Until recently, only the first approach has been extensively employed, for fairly evident reasons.

Analysis is always the first step in science. Carrel and Baker sought to analyze embryo juice by a variety of physical and chemical means. They concluded that animal fibroblasts could not utilize simple compounds such as amino acids but required complex proteoses or polypeptides and, in addition, unknown trephones, alexins, and other substances. Their work never led to the elimination of either plasma or serum from the nutrient. Simms has approached the problem in a somewhat similar fashion by the use of his "ultrafiltrate" which, while it is probably proteose in nature, is still an unknown of considerable complexity.

It is probably true that, as recently as ten years ago, no other approach to the problem offered any considerable hope of success. For, by then, our knowledge of the nutrition of organisms had hardly passed beyond the carbohydrate-protein-fat stage. We had some glimpses of the significance of vitamins and hormones, but no usefully precise knowledge. That, however, is no longer true. We now know the vitamin requirements of a great many organisms, and we have a pretty good picture of the amino acids and other protein constituents required. Moreover, collation of our information on a wide variety of organisms permits us to make certain generalizations which can be extended to organisms not yet studied. When we find thiamin an essential constituent in the diet of chicks, fungi, and tomato roots, we are safe in assuming its necessity for fibroblasts, since fibroblasts are, themselves, "elementary organisms". The fact that some organisms can make their own thiamin at varying levels of sufficiency, does not at all affect the conclusion that we should try thiamin as a constituent of any nutrient which we wish to devise. When we find histidine an essential constituent in the diets of rats and diphtheria bacilli, or glycine an effective nutrient for tomato roots and also an important constituent of anti-shock synthetic serum, we can look upon them as probable constituents of tissue nutrient solutions, even though histidine is not essential for maintenance of nitrogen balance in man, and glycine can be dispensed with by rats. When we find biotin cropping up as the "bios" of yeast, as a co-enzyme in legume bacterial respiration, as a nerve vitamin, and as an essential nutrilite for *Neurospora,* we can look upon it as a *general* requirement. This means that we have more than the beginning, I should say, on which to build a real synthetic nutrient for animal tissue cultures and thus establish the necessary base line for a real study of tissue nutrition.

Has this information been made use of adequately? I think not. Baker, and Baker and Ebeling have introduced certain vitamins into the nutrients for fibroblasts, but omission of pyridoxine, niacin, pantothenic acid, inositol, biotin, choline, and folic acid from any mention in their studies surely leaves much to be desired. Parker tells me that an amino acid mixture was used in some of Carrel and Lindberg's studies, but this was not published. Gey concluded that most amino acids were toxic, as one would gather from the literature of Carrel, Baker, Burrows, and others. While I feel fairly sure that others have thought along these same lines and must have carried out some investigations, I am familiar with only three groups which appear to have attacked the problem consistently and with some success. I refer to the groups centering around Simms in New York, Fischer in Copenhagen, and myself in Philadelphia.

I shall not attempt to speak for Dr. Simms. His work on the serum ultrafiltrate is an attempt to push back the bounds of ignorance by analytical methods and then to pile synthesis onto this base. His work on phosphate level as a factor in cell type dominance is important. Filtration, however, still does not, to my mind, establish a sufficiently clear base line, nor has Simms' group made adequate use of our newer nutritional knowledge. I think that much can be done with ultrafiltrate as a base, much more than has been done, by way of superimposing various growth-requirements thereon. But, at best, these will constitute a partial, though important, answer.

Dr. Fischer's work in Copenhagen has been pushed a little farther in the direction I have in mind. He has dialyzed his plasma to a point where it will no longer support any growth, and then added synthetic solutions thereto. With the nutrient constituents of plasma especially in mind, Dr. Fischer has tried an amino acid mixture equivalent to fibrin, with very great success. But he has not made extensive tests of the other nutrient requirements and has now reverted to the analytical approach. His, too, is therefore only a partial answer.

The lack of adequate consideration of these factors has long impressed me and, three years ago, I undertook to try my hand at remedying the situation, not without hesitation, for my training as a plant physiologist is rather far removed from oncological physiology. It has, on the other hand, given me the occasion to acquaint myself with a broader range of the nutrition literature than, I think, is true of most tissue culturists. In planning this work, I have considered it desirable not to stop at the level of an ultrafiltrate or a dialyzed plasma, but to go all the way to a completely synthetic medium, eliminating both plasma and embryo juice, growing my tissues in test tubes directly on glass, without any organic framework, and substituting a liquid medium of known constitution for the usual brew. Not all tissues will attach themselves directly to glass, but chick embryo heart muscle, skeletal muscle, and epithelium will all establish a satisfactory attachment if allowed to fix themselves for about two hours in sealed tubes before adding the liquid phase.

TABLE 1

NUTRIENT USED IN CULTIVATION OF CHICK TISSUES *In Vitro*

Inorganic factor	mg./100 ml.	Vitamin	mg./100 ml.
dextrose	850.0	thiamin	0.01
NaCl	700.0	pyridoxine	0.05
KCl	37.5	niacin	0.05
$Ca(NO_3)_2 \cdot H_2O$	21.0	riboflavin	0.01
$MgSO_4$	27.5	inositol	0.05
$Na_2HPO_4 \cdot 12H_2O$	14.5	ca-pantothenate	0.01
KH_2PO_4	2.6	biotin	0.01
$NaHCO_3$	55.0	choline	0 50
$Fe(NO_3)_3 \cdot 9H_2O$	0.14	folic acid	0.0001
dl-lysine HCl	15.6	carotene	0.01
dl-methionine	13.0	vitamin A	0.01
dl-threonine	13.0	ascorbic acid	0.05
dl-valine	13.0	glutathione	0.10
l-arginine HCl	7.8	cysteine HCl	0.10
l-histidine HCl	2.6	phenol red	1.50
dl-isoleucine	10.4		
dl-phenylalanine	5.0		
l-leucine	15.6		
dl-tryptophane	4.0		

The nutrient developed for this work (TABLE 1) differs, in many respects, from all others in the literature. The inorganic portion replaces $CaCl_2$ with $Ca(NO_3)_2$, and $MgCl_2$ with $MgSO_4$. It adds $Fe(NO_3)_3$ as a source of iron. The dextrose concentration is increased from 0.1 per cent to 0.8 per cent. There have been other minor concentration changes. Vitamins supplied are thiamin (B_1), riboflavin (B_2), pyridoxine (B_6), niacin, biotin, calcium pantothenate, ascorbic acid (C), carotene, inositol, choline, and folic acid. The necessity for all these has not been individually determined: this is, admittedly, a shot-gun mixture which will have to be studied in more detail, which can probably be simplified, and may ultimately need to be enlarged. All we can say, at present, is that chick embryo tissues do better with all of these vitamins included than they do when all are omitted.

Perhaps the most clearly important modification has been, contrary to the previous literature on the subject, the inclusion of amino acids in the nutrient. Those added for their nutritive value were chosen to correspond with Madden and Whipple's blood plasma substitute, and included lysine, methionine, threonine, valine, arginine, histidine, isoleucine, phenylalanine, leucine, and tryptophane. Cysteine and glutathione (glycine, glutamic acid, cystine) were added as buffers against breakdown of the vitamin A solution. Addition of this amino acid mixture increases survival of chick embryo tissues from the control level of five or six days to more than 50 days. That the amino acids are of crucial importance for growth of tissues and that, contrary to the earlier belief, they *can* be utilized in this form is, I think, now established beyond question by my own work and that of Fischer. Here, again, the mixture has not yet been subjected to detailed study and can, in all probability, be improved by simplification and expansion.

Fischer tried this same amino acid solution superimposed on a plasma base. He concluded that it was without effect, but that a somewhat different solution, based not on Madden and Whipple but on Bergmann and Niemann's fibrin analyses, was an effective nutrient. I have not yet tried Bergmann and Niemann's formula. I think it may be significant, however, that it contains both cystine and glutamic acid, which were not present in Madden and Whipple's mixture as supplied by Fischer, but were supplied in my work as glutathione in the vitamin mixture. Moreover, Fischer's tests were made without vitamin supply and, presumably, without a known iron supply. It will be interesting to compare the two solutions under my conditions.

Using this admittedly complex and insufficiently elucidated but nevertheless fully synthetic nutrient, it has been possible to maintain chick embryo heart muscle beating for six weeks with repeated renewal of nutrient; skeletal fibroblasts in good condition for eight weeks; and epithelium and neural elements for somewhat shorter periods. The nutrient has not been tested extensively on adult avian tissues or mammalian tissues. It has not been tested at all on malignant tissues. All this represents a beginning, hardly more.

While the subject of this paper was to be nutrition of malignant tissues, they have hardly been mentioned so far. Lambert and Hanes 35 years ago, Fischer 20 years later, and others at various times, have shown that malignant tissues tend to liquefy plasma somewhat more rapidly than do normal tissues. As a concomitant of this liquefaction, which appears to be a digestive process, malignant tissues are able to survive and grow at the expense of substrata which are insufficient for normal tissues. Carrel and Baker studied the effects of various tissue digests on sarcoma cells in 1928, while Demuth and von Riesen, at about the same time, carried out similar studies with carcinoma. Fischer and his co-workers also made studies of the same sort. The conclusion has been more or less identical in all these studies: that malignant tissues possess an enzymatic armamentarium which permits them to survive at the expense of nutrient conditions which are not suitable for normal tissues. It has been assumed that these conditions represent simpler compounds, that malignant tissues can use amino acids while normal tissues require proteoses. However, this was an assumption based on indirect evidence and, as Fischer and I have recently shown, it was an unsound assumption, for normal tissues can *also* utilize simple compounds. There obviously exists the alternative possibility that malignant tissues release amino acids from plasma, but that, in doing so, they are merely duplicating the quantitative conditions which already exist in embryo juice. Fischer's demonstration that normal tissues will grow, though slowly, on plasma alone, but will not do so on dialyzed plasma, supports this view, adding one more case in which the difference between normal and malignant tissue appears to be quantitative and not qualitative.

This leaves us still with very little in the way of facts concerning the nutrition of malignant tissues. I do not like to deal in theories but, in the absence of facts, I am prepared to run the risk of entering into a short discussion of the possible nature of malignant tissues, drawing upon my background in plant physiology for analogies.

We possess, in plants, certain tumors called grown galls which bear many resemblances to animal cancers and can be manipulated in many ways not yet available to animal oncologists. Braun's work on heat therapy in crown,gall tumor production on *Vinca rosea* has established that the irreversible change from normal to malignant cells goes to completion in not more than 72 hours and is probably complete in considerably less time than that. Earle has, in a somewhat more tentative way, shown that the induction of malignancy in cultures of mouse tissues by use of methylcholanthrene is, likewise, complete in not more than six days. Black's work on the involvement of wounding in the induction of virus tumors in plants, and Rous and Kidd's studies on the effects of wounds and tar in accelerating the malignant change of the Shope papilloma, also indicate that the actual cancerization may be a relatively abrupt process.

The mutation theory of cancer is an old one. Dr. Tatum discussed its merits in one of the preceding papers, and some of my remarks may overlap his. , The mutation theory has, in general, been rejected because, unlike all other known mutations, it seems to be consistently unidirectional. You either have cancer or you have nothing. Mutations just don't act that way. But is this a true statement of the facts? Are we not setting up our conditions in such a way that this is the only possible *visible* result? In plant tumors, we have reason to believe that one of the major factors in the development of malignancy is the sudden enhancement of the cell's capacity to synthesize the auxines necessary for its independent growth. Normal plant cells do not make quantities of these substances sufficient for autonomous growth, while malignant cells do. Now, where else in plant biology do we find a similar sudden change in the synthetic capacities of a cell? You will realize, immediately, that it is in the biology of *Neurospora* as elucidated by Beadle and Tatum. Here, mutations have occurred which result in sudden changes in the metabolism of the cell. These mutations are lethal, except when segregated by careful nutritional studies, based first on the establishment of a synthetic basic nutrient. The "wild-type" nutrient contains only biotin as accessory nutrilite. For the mutants, single substances or simple substance combinations must be added thereto. Now, what would happen if similar mutations were to occur in the somatic tissues of a mammal? These tissues are dependent on the blood stream for such of their needs as they are not capable of covering by their own metabolism. We know that this dependence is great. If a mutation occurred by which the normal capacity of a particular cell to synthesize a substance supplied by the blood stream, such as pyridoxine, was lost, that loss would be without any visible effect. If a mutation occurred by which the synthesis of a substance required for survival and not supplied by the blood stream were lost, the cell would die and a minute necrotic locus would result which would soon be repaired. If a mutation occurred by which the capacity to synthesize a substance such as insulin (which is supplied by the blood stream) was *acquired,* this would result in a notable response only if the concentration of that substance normally present was a limiting factor in tissue growth. Thus, knowing that thiamin is normally present in the blood stream at considerable concentration, and that thiamin is one of those substances of which one cannot get an overdose, it is clear that a mutation by which the cell's capacity to synthesize thiamin was either reduced or enhanced would produce no visible result whatever. Such a mutation would never be detected by our present methods. The same would be true of any substance which, while not present in the blood stream, is nevertheless neither necessary nor injurious to the cell. Only if the substance was a limiting one, would the response be visible. Now, in plant tumors, we have such a response. Plant tissues require auxine for autonomous growth. Their synthetic capacity furnishes a limiting condition in their behavior. When a capacity to synthesize auxine at a high level is acquired, there is an immediate escape from dependence on the surroundings, in the form of cancerization. There is reason to believe that the steroids may play a role in carcinogenesis in animals. They are probably synthesized at low levels by many cells, but synthesis at a level significant for growth is ordinarily restricted to certain cells whose metabolism is adapted to handling them without untoward results. If, now, a tissue cell should abruptly acquire the capacity to synthesize steroids at a level which would free it of dependence on its surroundings, this might serve as the trigger for the development of what we call malignancy. I do not wish to emphasize this particular example. What I do want to stress is that, with our present methods of diagnosis, *this is the only kind of a mutation that we would recognize.* This is not quite true, for the abnormal production of melanin *is* recognizable in melanotic moles which, as we well know, are often associated with or precursors of malignancy. Alcaptanuria, phenylpyruvic idiocy, albinism, etc., are all examples, in man, of metabolic somatic mutations which *are* recognizable. It seems to me extremely likely that similar mutations are occurring in the body constantly, and that they are, for the most part, unknown to us because, while melanosis, tumorization, and malignancy are visible, others are beyond our methods of diagnosis.

If, on the other hand, we had means of growing cells *in vitro* under conditions which could be adequately controlled, there would immediately become available to us the whole gamut of nutritional, respirational, enzymatic, and other studies which are used in diagnosing mutations in *Neurospora* and which have, in the past decade, given us such a large measure of command over the metabolic destinies of man and beast. It is for that reason that I have set up this hypothesis in discussing the nutrition of malignant tissues *in vitro*, for it is a hypothesis which can only be examined by such an approach. There seems to me to be no doubt that such an approach is needed and, although, as I must emphasize again, I can at the moment contribute little in the way of factual knowledge to the problem of malignant tissue metabolism, I hope to have pointed the way to possible methods of remedying this lacuna.

REFERENCES

1. **Baker, L. E., & A. Ebeling**
   1939. Artificial maintenance media for cell and organ cultivation. I. The cultivation of fibroblasts in artificial and serumless media. J. Exp. Med. **69**: 365–378.
2. **Bergmann, M., & C. Niemann**
   1936. On blood fibrin, a contribution to the problem of protein structure. J. Biol. Chem. **115**: 77–85.
3. **Fischer, A.**
   1941. Die Bedeutung der Aminosaeuren fuer die Gewebezellen in *vitro*. Acta Physiol. Scand. **2**: 145–188.
4. **Madden, S. C., *et al.***
   1943. Ten amino-acids essential for plasma protein production effective orally or intravenously. J. Exp. Med. **77**: 277–295.
5. **White, P. R.**
   1946. Cultivation of animal tissues *in vitro* in nutrients of precisely known constitution. Growth **10**: 231–289.

EFFECTS ON THE GENESIS AND GROWTH OF TUMORS ASSOCIATED WITH VITAMIN INTAKE

By HAROLD P. MORRIS

National Cancer Institute, National Institute of Health, United States Public Health Service, Bethesda, Maryland

Nutritional factors, such as vitamins, which may be involved in the genesis of cancer should be clearly distinguished from nutritional constituents which influence the growth of a tumor that is already present, because the two processes may be quite different. Likewise, it is not possible to make extensive generalizations from the effects obtained on one tumor to those which may prevail for other types of tumors. It is exceedingly difficult as well as hazardous, therefore, at the present time, fully to evaluate the role of vitamins in the origin and growth of cancer. It is, nevertheless, reasonable to believe that certain essential dietary constituents, including vitamins, are involved in the genesis or growth of tumors.

The experiments described in this paper were conceived and carried out to show the effects of different vitamin levels on the genesis as well as the development of tumors. In the change to malignancy, the requirements of the altered cell may be such that it now requires more, less, or none of some vitamin essential to the normal cell. Two objectives of these nutrition studies have thus been the determination of the influence of vitamin intake (1) on the *in vivo* growth of tumors already formed, and (2) on the genesis of cancer. The availability in pure form of pantothenic acid, pyridoxine, thiamine, and riboflavin made feasible the initiation of such experiments.

The problem so far has included investigations of the following points: (1) the quantitative appraisal of the host's requirement for the given vitamin; (2) effect of a rapidly produced, extreme deficiency of the vitamins on the growth of tumors; (3) effect on the growth of the tumor of a vitamin intake just adequate to maintain body weight of the non-tumor-bearing adult animal; (4) other effects of vitamin deficiency or supplementation on tumor and host; and (5) influence of a prolonged partial vitamin depletion of the animal on the genesis of spontaneous mammary tumors.

Not all of the above points will be fully presented in this discussion, and many details already published or to be published will be omitted in order to give briefly an overall picture of the *in vivo* studies so far completed.

MATERIALS AND METHODS

The diets used in our experiments contained all the dietary essentials except for the single one being studied. Two types of short-term deficiency experiments were carried out: (1) Rapid extreme depletion of the animal in a given vitamin by feeding the animal a diet adequate in all other dietary essentials except the one under investigation, with the diet as completely devoid of that vitamin as it was possible to prepare it. (2) Partial depletion of the animal in a given vitamin by adding just sufficient of the vitamin to the deficient diet to permit maintenance of body weight of the non-tumor-bearing animal.

In producing the first type of short-term deficiency, a very rapid and extreme depletion of the animal occurred during a period of a few weeks. In the second type, the animal was never sufficiently depleted to develop clinical symptoms of deficiency.

Long-term partial depletion of riboflavin was produced in mice by allowing the animal to consume *ad libitum* a diet which contained just sufficient riboflavin over a prolonged period to prevent the development of gross clinical symptoms of ariboflavinosis.

The requirements for pantothenic acid, riboflavin, pyridoxine, and thiamine were carefully determined for growth of young mice and maintenance of adult body weight of animals without tumors prior to the experiments on adult mice with tumors.[1, 4, 6] Strain C3H mice were used throughout. The spontaneous mammary tumor arising in this strain was utilized for the studies on tumor genesis and growth. Growth rate of the tumor was determined at weekly intervals over a period of several weeks. The size of the tumors was determined by measuring two diameters of the tumor with a caliper and expressing the size as the product of these two measurements.

EXPERIMENTAL RESULTS

Growth of Tumors During Short-Term Extreme and Partial Vitamin Deficiencies

PANTOTHENIC ACID. Short-term severe pantothenic acid deficiency inhibited growth of normal young mice and resulted in loss of body weight of adult animals.[2] The influence of a deficiency of pantothenic acid on tumor growth was determined by placing C3H mice bearing small spontaneous mammary adenocarcinomata on this diet, which was extremely deficient in pantothenic acid.[3] The average rate of growth of tumors was followed for several weeks. During this time, the average growth rate was definitely retarded, as shown by curve *B* of FIGURE 1. The much more rapid average rate of growth of the tumors in control animals whose diet was supplemented with a large excess of pantothenic acid, is shown by curve *A* of FIGURE 1. The supplementation of the diet with

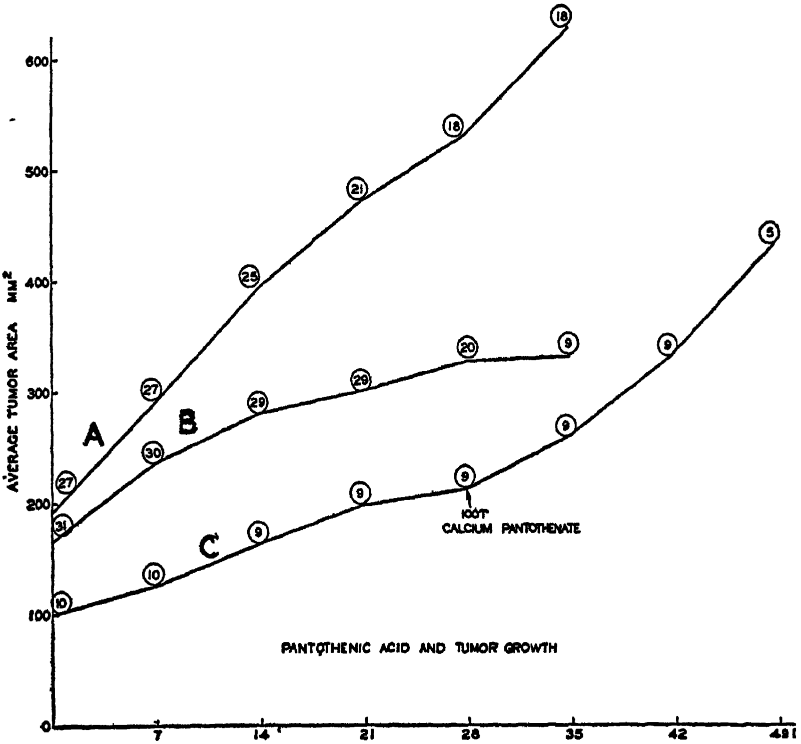


FIGURE 1. Effect of pantothenic acid on tumor growth rate. The numbers in the circles represent the number of tumors used to obtain the average values on the curves.

pantothenic acid after four weeks of extreme deficiency in a third group of animals also increased the average rate of growth of tumors, as shown by curve *C* of FIGURE 1. It was concluded, from these experiments, that pantothenic acid deficiency retarded the growth of the mammary tumor.

PYRIDOXINE. This vitamin is essential for the adult mouse. Rapid loss in body weight and death in approximately seven weeks occurred in adult mice fed, in a short-term experiment, a diet extremely deficient in pyridoxine,[4] as shown by curve *A* of FIGURE 2. The addition of pyridoxine to the deficient basal diet, beginning with the fifth week, brought about a very rapid recovery of the body weight (curve *B*, FIGURE 2). No similar loss in body weight was found with tumor-bearing mice given the same pyridoxine-deficient diet.[4] In a second group of tumor-bearing mice, after three weeks of deficiency only a slight increase in body weight was noted following the addition of pyridoxine to the diet, as shown in FIGURE 3B.

The amount of pyridoxine in the diet did not affect the rate of growth of the spontaneous tumor. The average rate of growth of the tumors during a 6-week period of deficiency, as shown in FIGURE 3A (solid line), was the same as the rate of growth after three weeks of deficiency followed by pyridoxine supplementation for five weeks (solid and broken line).

THIAMINE. One of the first effects observed in adult mice deprived of thiamine was a self-induced restriction in food consumption[4] which rapidly resulted in complete failure to ingest food even though the period of depletion was relatively short. In the absence of thiamine, Morris and Dubnik[4] observed a rapid loss in body weight of adult mice, followed by death in 3-4 weeks. That death was brought about primarily through this failure to ingest food rather than from a deficiency in thiamine, was shown by another group of mice receiving the thiamine but restricted to the same quantity of food consumed voluntarily by the thiamine-deficient animals. In this diet-restricted group, death occurred at approximately the same time as in the deficient group.

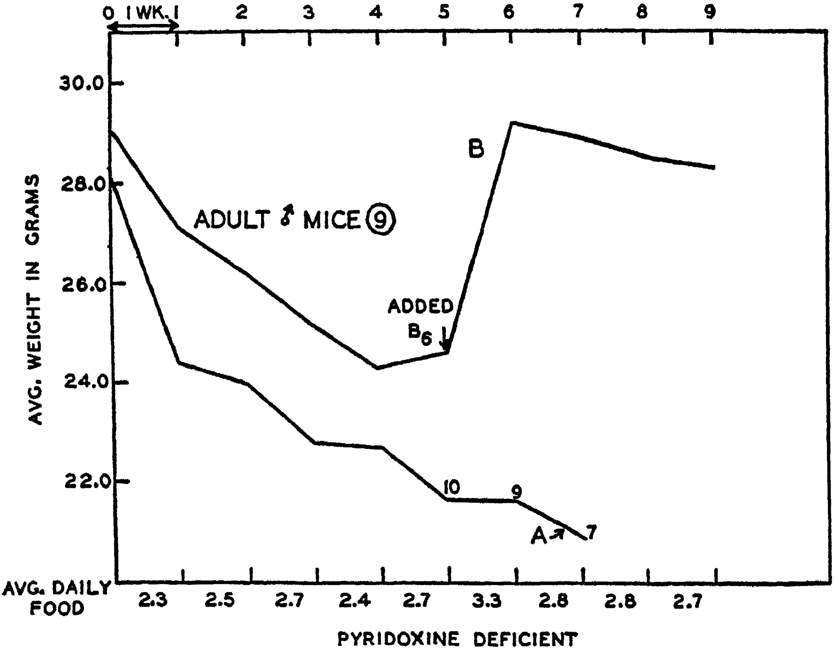


FIGURE 2. Average body weight of adult C3H mice fed a pyridoxine-deficient diet, Curve *A*. Curve *B*, avg. body weight of mice deficient for 5 weeks, then supplemented with pyridoxine. The average daily food intake is given by the figures at the bottom of the chart.

PYRIDOXINE EXPERIMENTS --TUMOR MICE

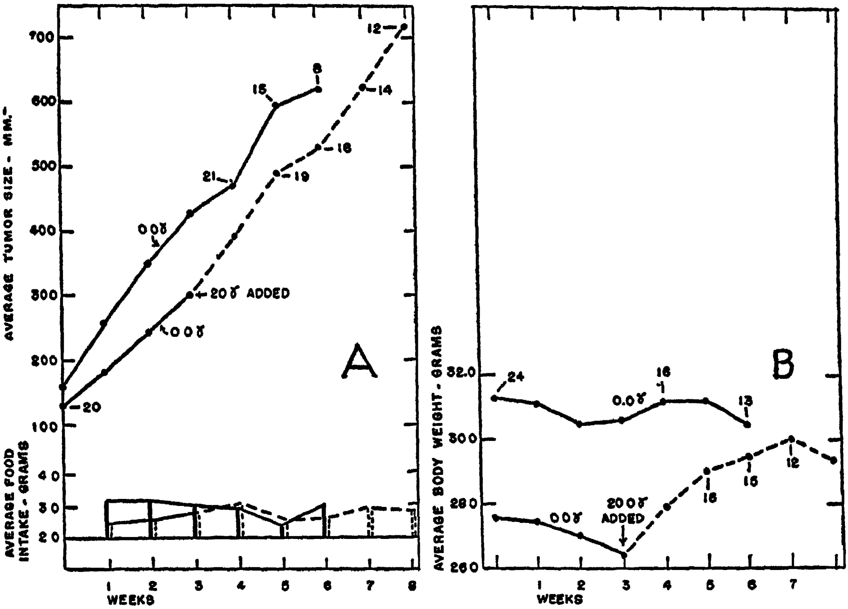


Figure 3. B. Average body weight of adult tumor-bearing C3H mice fed a pyridoxine-deficient diet (solid line), and after supplementation with pyridoxine (broken line).
A. Average rate of tumor growth in mice fed a pyridoxine-deficient diet for five or more weeks (solid line); same for mice fed the deficient diet 3 weeks and then supplemented with pyridoxine (solid-broken line). The numbers in both A and B represent the number of animals or tumors used to determine the point on the curve. At the bottom of figure A, the average amount of food consumed by the tumor-bearing mice is given.

That the retarded rate of growth of the spontaneous mammary tumor in mice extremely deficient in thiamine was also due primarily to the failure of the animal to ingest an adequate quantity of food, was demonstrated in analogous experiments carried out with tumor-bearing animals. The rate of growth of tumors on the thiamine-deficient diet is illustrated by Figure 4, right half, curve *D*, while curve *E* shows the growth obtained when the amount of thiamine-supplemented diet was restricted to the same quantity of food consumed voluntarily by the thiamine-deficient animals. The growth rate of the tumors in both groups was essentially parallel.

The average growth rate of the tumor was not affected in mice ingesting diets containing 1-2 γ of the vitamin per gram of food, an amount just sufficient to maintain body weight of adult tumor- and non-tumor-bearing mice (Figure 4, right half, curves *A*, *B*, and *C*).

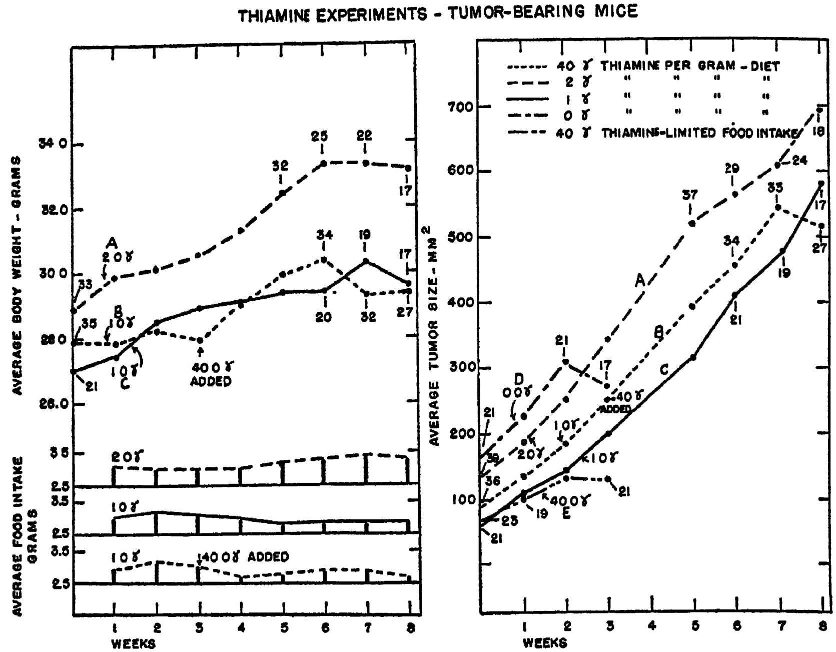


FIGURE 4. Average body weight of tumor-bearing C3H mice fed different amounts of thiamine: curve *A*, 2 γ per gram of food; curve *B*, 1.0 γ for 3 weeks, then 40 γ; curve *C*, 1 γ; average rate of tumor growth on different levels of thiamine supplementation as indicated by the legend. Mice shown in curves *D* and *E* received same amount of food. The numbers at different points on the curves refer to the number of animals used to obtain the average values.

A method to circumvent the voluntary restriction of food intake of thiamine-deficient mice was devised.[4] This was accomplished by forcibly feeding two groups of paired mice the same quantity of thiamine-deficient or thiamine-supplemented food, respectively. The body weight was reasonably well maintained in both groups by this procedure (upper two curves of FIGURE 5). The average growth rate of the tumors in thiamine-deficient, forcibly fed mice appears to be slightly accelerated (FIGURE 5, curve 110), when compared to the average rate of growth of the tumors in the thiamine-supplemented, forcibly fed group (FIGURE 5, curve 79).

The slightly decreased average rate of growth of the tumors in the thiamine-supplemented, forcibly fed group may even indicate a depressing effect of thiamine on the growth of the tumor. Dobrovolskaïa-Zavakskaïa[5] has reported retarded growth of the mouse spontaneous mammary tumor after daily, subcutaneous injections of 40-60 γ of thiamine, amounts similar to the amounts ingested in the above experiments. Such injections were also reported to prolong the life of the animal.

RIBOFLAVIN. Extreme deficiency of riboflavin in non-tumor-bearing adult mice was shown by Morris and Robertson[6] to result in loss of body weight, followed by death in about nine weeks. Young mice failed to grow after the initial stores of riboflavin were used up, maintained their body weight for several weeks, and then died from the effects of riboflavin deficiency, with only slight loss in body weight.

Morris and Robertson[6] found that extreme deficiency of riboflavin decreased the rate of tumor growth, when compared to the average rate of growth of the tumor in mice on a riboflavin-supplemented diet. The tumors of mice extremely deficient in riboflavin showed a gradually decreasing average rate of growth which was more pronounced during the terminal period of the experiment (FIGURE 6, curves *A* and *B*).

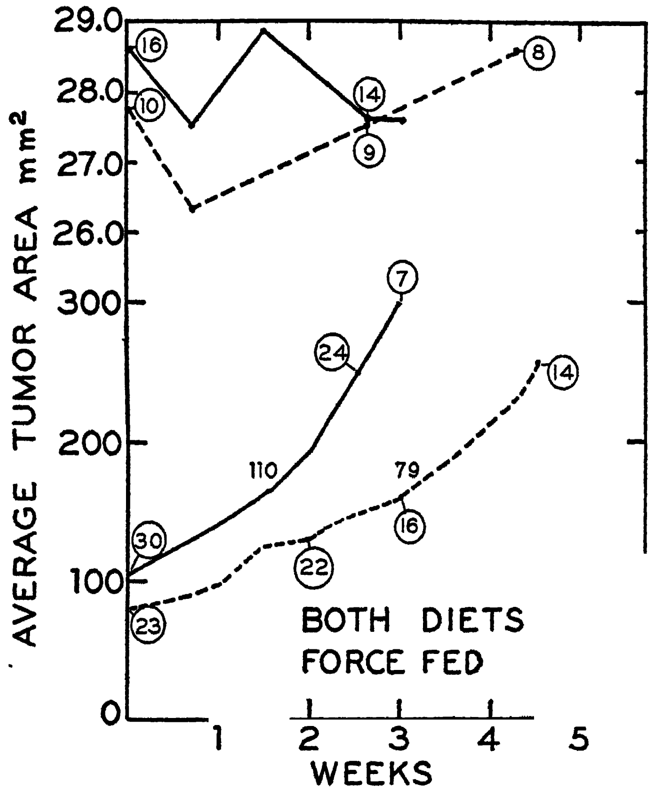


FIGURE 5. The average rate of tumor growth (lower two curves), and average body weight (upper two curves), of mice forcibly fed a thiamine-deficient diet (solid lines), and a thiamine-supplemented diet (broken lines). The circled numbers represent the number of animals used to obtain the averages.

Riboflavin supplementation in another group of mice after three weeks of deficiency caused a striking stimulation of the depressed growth obtained during the deficient period (FIGURE 6, curve *C*). Gross clinical symptoms of riboflavin deficiency were just beginning to appear in the host at the time riboflavin was added to the diet. A fourth group fed the deficient diet for 3 weeks, then given a riboflavin-supplemented diet for 2 weeks, followed by a second period of riboflavin deficiency, showed depression of tumor growth during both deficient periods and stimulation during riboflavin ingestion as shown by FIGURE 6, curve *D*.

No decrease in the average rate of growth of the tumor was observed in mice partially depleted in riboflavin,[6] or if the level of riboflavin given was just sufficient to maintain the body weight of the non-tumor-bearing adult mouse. The average rate of tumor growth in mice on this maintenance level of riboflavin is illustrated by the tumor growth curve shown in FIGURE 7, curve *E*. Mice on this maintenance level of riboflavin for four weeks were then given a diet supplemented with a large excess, 10 γ per gram of food, of riboflavin. These animals did not show any stimulation in the growth of their tumors after the addition of riboflavin. This was in contrast to what was observed in groups made extremely deficient in riboflavin prior to supplementation. The decreased rate of growth of tumors in mice extremely deficient in riboflavin is shown for comparison (curve *B* of FIGURE 7, reproduced from FIGURE 6).

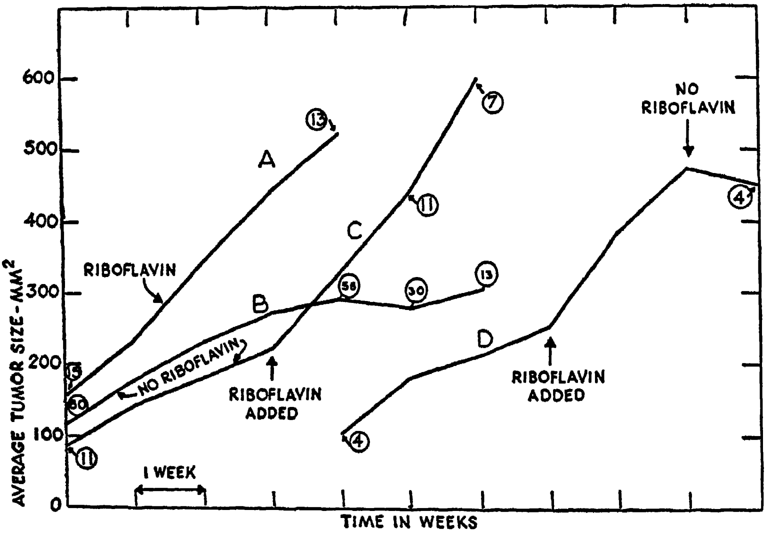


FIGURE 6. Average rate of tumor growth in acute riboflavin deficiency. *A*—riboflavin-supplemented; *B*—deficient; *C*—deficient 3 weeks, then supplemented; *D*—deficient 3 weeks, supplemented 2 weeks, then deficient 1 week. The circled numbers are the number of tumors used to obtain the average values.

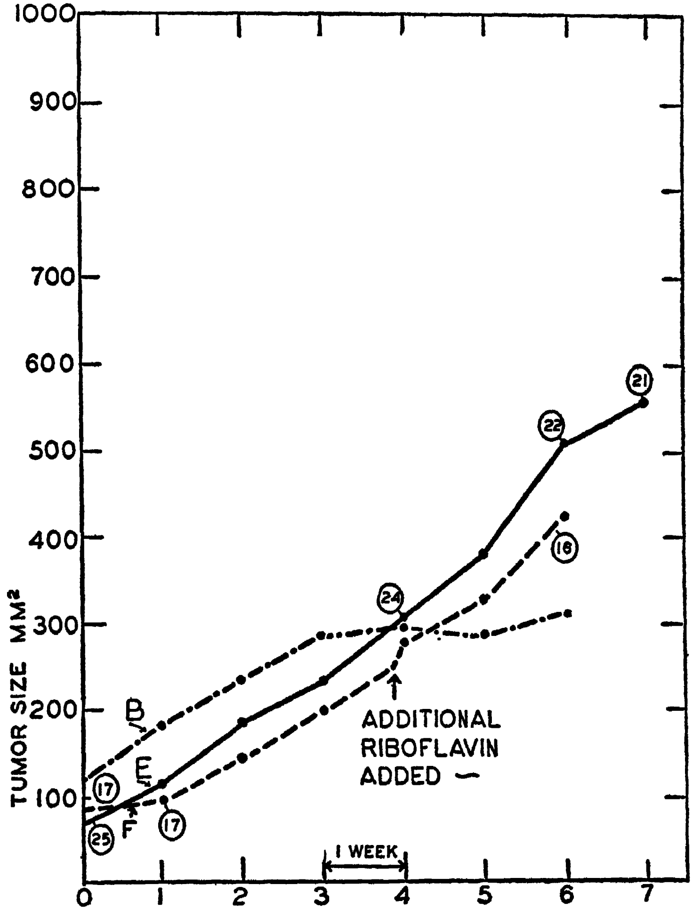


FIGURE 7. Average growth rate of mammary tumors: *B*, riboflavin-deficient; *E*, supplemented with a maintenance level of riboflavin; *F*, supplemented with a maintenance level of riboflavin. After 4 weeks, increased to 10 γ riboflavin per gram of food.

It would appear from these data that the host must be made extremely deficient in riboflavin before noticeable retardation of the growth of the spontaneous mammary tumor occurs.

Other Effects of Riboflavin Deficiency or Supplementation on Tumor and Host

TISSUE CONCENTRATION. Additional information was sought on the role of riboflavin in the growth of tumors by studying the riboflavin concentration of tumor, liver, and muscle, three tissues likely to contain large amounts of this vitamin.[6] A decrease in riboflavin concentration approximating 50 per cent of normal in both tumor and muscle was found in animals after ingestion of the basal diet deficient in riboflavin for six weeks, as shown by the data given in TABLE 1. The riboflavin concentration in the liver, however, had decreased by 50 per cent in four weeks with an insignificant further decrease by six weeks, as shown in TABLE 1. The decrease in riboflavin concentration found in the liver followed rather closely the observed decrease in body weight of the deficient mice, while the somewhat slower decreasing concentration noted in the tumor of the deficient mice seemed to follow more closely the decreased rate of tumor growth observed in these deficient animals.

TABLE 1

AVERAGE RIBOFLAVIN CONCENTRATION IN MOUSE TISSUE

Mice	Period of deficiency, weeks	Mice analyzed, No.	Diet No.	Riboflavin per gm. of dry tissue Liver μg.	Muscle μg.	Tumor μg.
Normal	0	5	67	115.6	13.8	
	6	5	66	58.5	9.1	
Tumor-bearing	0	5	Stock*	112.8	16.1	29.9†, ‡
	0	5	67	120.7	16.3	29.5
	1	5	66	100.0	12.4	29.5‡
	2	5	66	99.2	12.4	25.0
	3	5	66	81.5	11.6	26.5
	4	6	66	61.8	10.3	19.4
	6	7	66	64.6	11.1	16.5
Standard deviation				7.75	1.18	1.09
Standard error				4.90	.75	.69
Amount necessary for significant difference at 5 per cent point				9.97	1.50	1.46

* Purina dog chow.

† Pooled sample, calculated from fresh-weight analysis.

‡ Not used in statistical calculations because the analysis was determined from a pooled sample.

A comparison of the decreasing rate of growth of the tumor in riboflavin-deficient mice (FIGURE 8, curve *B*) shows a striking similarity to the riboflavin concentration of the tumor in riboflavin-deficient mice (curve *C*). The data for curve *C* were obtained from TABLE 1. The greatest decrease in both the rate of growth and the riboflavin content of the tumors occurred during the third to fifth weeks of the vitamin deficiency.

TOTAL SIZE OF TUMORS. All palpable tumors arising in two different groups of mice receiving riboflavin-deficient and supplemented diets were combined, and the average total size of tumors per mouse calculated for each group. The mice were started on the diets when the first tumor was observed. The supplemented group received 10 γ of riboflavin per gram of food. The deficient group received the basal diet deficient in riboflavin.[6]

A striking decrease was noted in the average total size of the tumors in the deficient basal diet group, when compared to the size of tumors in the group on the diet supplemented with riboflavin. This is illustrated in curves *A* and *B* of FIGURE 8.

A striking decrease was found in the rate of growth of tumors in mice deficient in riboflavin compared to mice receiving a supplement of riboflavin, as noted in FIGURE 6, curves *A* and *B*. This difference in rate of

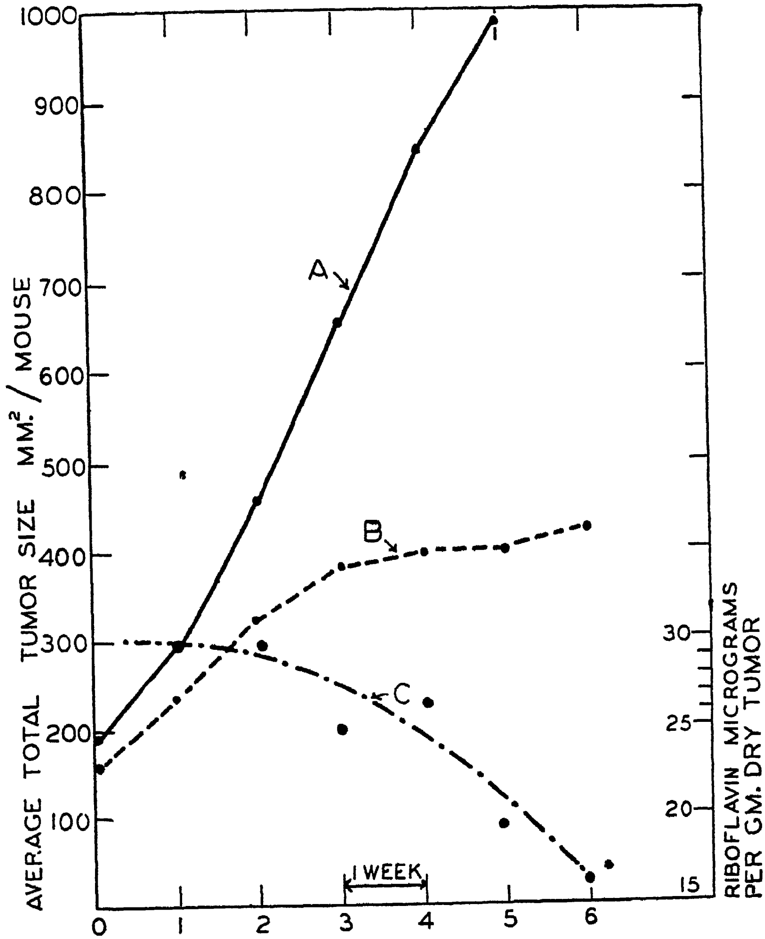


FIGURE 8. Average total tumor size and tumor riboflavin concentration. Curve *A*, riboflavin-supplemented mice; curve *B*, riboflavin-deficient mice; curve *C*, tumor concentration of riboflavin in deficient mice.

growth was obtained by using only those tumors present when the experiment was started, and did not include any of the additional tumors which developed during the course of the experiment. When all the tumors that arose in the two groups were used and compared (as illustrated in FIGURE 8, curves *A* and *B*), however, a much more striking difference was found between the supplemented and deficient groups. This additional effect is due to the increased number of tumors arising in mice receiving the riboflavin supplement.

TUMOR SIZE AND LENGTH OF SURVIVAL. The amount of riboflavin in the diet affected the length of survival of mice with mammary tumors.[6] This was a direct relation, for it was found that tumor-bearing mice, when fed the riboflavin-deficient diet, survived an average of 40 days, while the animals fed the supplemented diet (10 γ per gram of food) survived, on the average, 62 days. If, however, the mice were first depleted in the vitamin for three weeks and their diet then supplemented with the 10-γ level of riboflavin, their survival time was extended to 76 days.

Thus far, it has been shown that short-term extreme riboflavin deficiency (1) decreases the rate of tumor growth, (2) shortens the length of survival, and that (3) riboflavin supplementation increases the number of tumors. To test the possibility that tumor size was correlated with survival of the animal, a statistical analysis for such a correlation was made of five groups of mice.[6] A large series of animals were used in this study, representing (1) riboflavin-deficient; (2) supplemented; (3) maintenance throughout; (4) maintenance for three weeks, then supplemented; and (5) depleted three weeks, then supplemented. The results of these calculations are expressed (FIGURE 9) by the regression curves of survival in days over total tumor size. No significant deviation from zero correlation was indicated between maximum tumor size and survival of the mice in any of the groups.

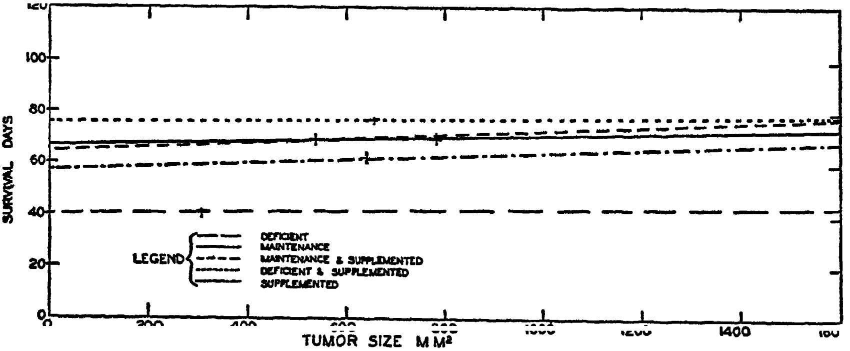


FIGURE 9. The linear regression curves of survival in days of mice after appearance of first tumor plotted against total size of the tumor at various levels of riboflavin intake, as indicated.

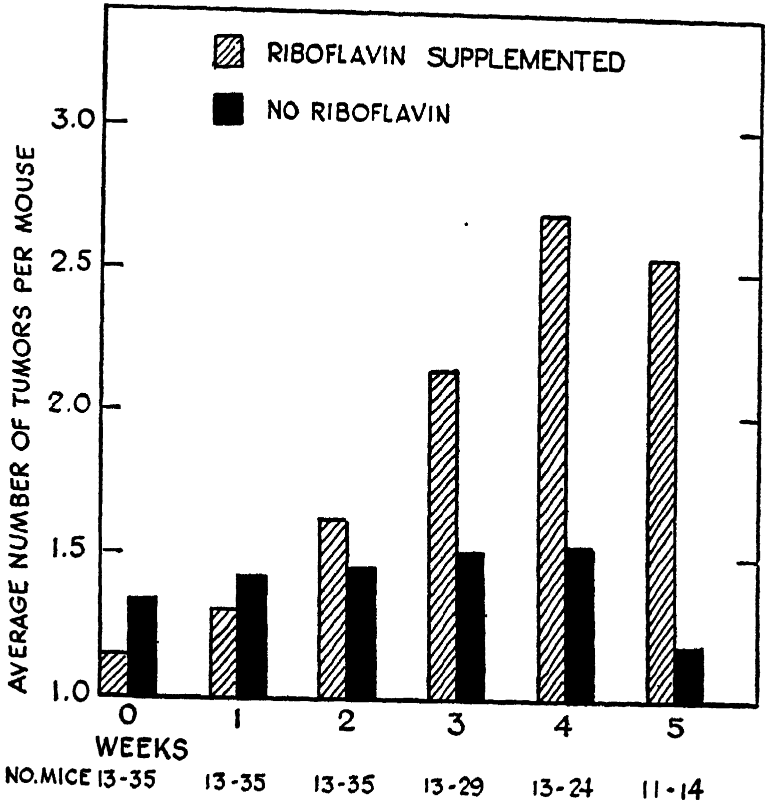


Figure 10. Average number of tumors per mouse in mice fed riboflavin-deficient and supplemented diets.

It would, therefore, seem justifiable to conclude that the size of the spontaneous mammary tumors of mice, once they become palpable, does not affect the length of time the animal will live, either on a riboflavin-deficient diet, a maintenance diet, or on one supplemented with an abundant amount of riboflavin. Furthermore, the survival of the tumor-bearing animal is not affected by the size of the tumor, but is directly related to the riboflavin supplementation or a combination of partial depletion followed by supplementation.

Riboflavin Intake and the Number of Tumors. It is generally recognized that female mice susceptible to the development of spontaneous mammary adenocarcinoma will develop more than a single gross tumor.[7] That the number of these tumors can be affected by dietary constituents, was first shown by Morris and Robertson[6] who found a direct relation between the amount of riboflavin in the diet and the number of gross tumors in C3H mice. They selected, at random, tumor-bearing mice ingesting a stock diet and placed them on experimental diets. They then compared, at weekly intervals, the number of tumors observed in one group of animals given the basal diet containing only 0.5 γ of riboflavin per gram of food with the number arising in the other group receiving this basal diet supplemented with 10 γ riboflavin per gram of food. There was no essential difference in the total quantity of food ingested by the two groups, but the group supplemented with riboflavin showed a striking increase in the number of tumors during a five-week period. The number of tumors in the deficient group increased slightly from 1.3 to 1.5 tumors per mouse, while, during the same period, the tumors in the supplemented group more than doubled in number by increasing from 1.2 to 2.7 tumors per animal. This striking difference between the two groups is illustrated in the bar diagram of FIGURE 10.

To see if the amount of riboflavin ingested for a prolonged period prior to tumor development also affected the number of tumors, graded amounts of riboflavin were given throughout the lifetime of the mouse.[4] Weanling female C3H mice were divided into three groups. Each group was fed a diet containing a different amount of riboflavin, and the experiment continued throughout life until death occurred following the development of mammary tumors. Eleven mice in one group were given, throughout their lifetime, a diet containing a minimum amount of riboflavin approximating 1.5 γ per gram of food. The animals of this group developed an average of 1.8 tumors per mouse and survived an average of 4.6 weeks following the appearance of the first tumor, as shown in column *A* of FIGURE 11. Forty-eight mice received about twice the minimum level of riboflavin, 3.3 γ per gram of food. These animals survived an average of 8.1 weeks after the appearance of the first tumor, and developed an average of 2.2 tumors per animal (column *B*, FIGURE 11). The third group of 40 mice received approximately seven times the minimum level of riboflavin, or 10 γ per gram of food. This group developed an average of 2.5 tumors per mouse during a survival period, after the first tumor appeared, of 6.4 weeks (column *C*). These data support the view that the effect of riboflavin on increasing the number of tumors occurs only after the appearance of the first palpable tumor, and that the effect is not due entirely to the length of survival of the animal after the first tumor develops.

Influence of a Prolonged Partial Riboflavin Depletion on Genesis of the Spontaneous Mammary Tumor

The foregoing experiments may also be interpreted as showing that decreasing amounts of dietary riboflavin resulted in fewer spontaneous mammary tumors. This suggests that the incidence of mammary tumors may be influenced by partial riboflavin deficiency, and offers a means of studying the influence of this vitamin on the genesis of mammary tumors

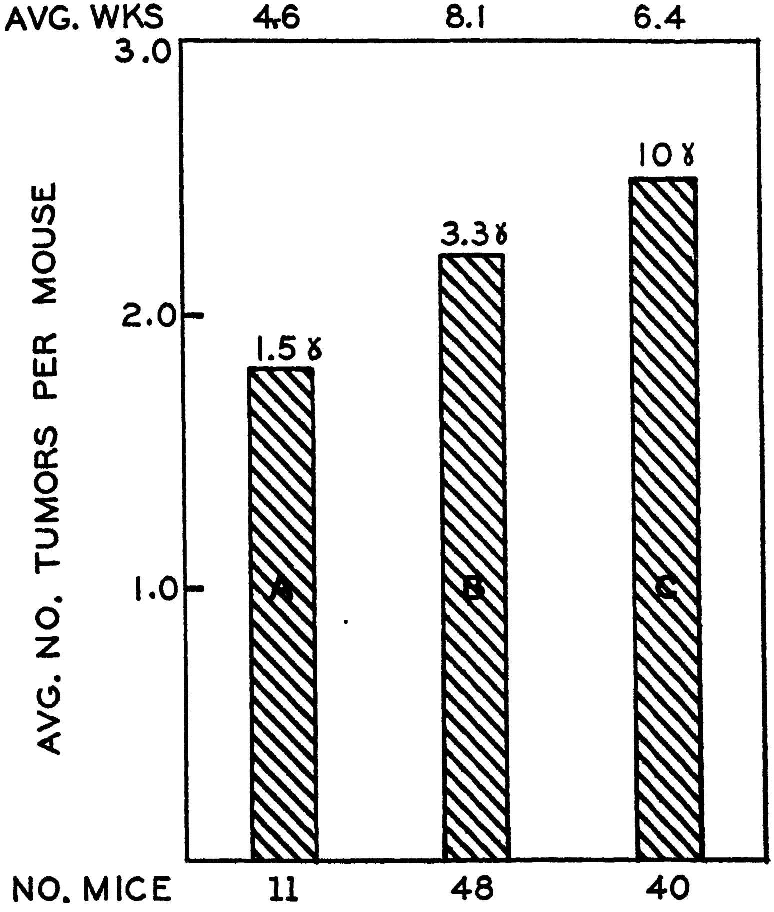


FIGURE 11. Average number of tumors per mouse in mice fed different amounts of riboflavin throughout life, expressed in micrograms per gram of food at top of columns. Values at top of figure represent average survival after appearance of the first tumor.

in mice. The mouse mammary cancer substrate, the mammary glands, must reach a certain stage of development before the mammary tumor will arise. Any dietary regimen, therefore, which influences tumor genesis should be initiated at an age prior to the full development of the mammary glands in order to obtain maximum effect.

Morris and Dubnik[4] studied the influence of partial ariboflavinosis on the spontaneous mammary tumor incidence in C3H virgin female mice. The experiments were initiated at weaning time and continued throughout the life of the animal. The animals were allowed just sufficient riboflavin to prevent severe and chronic clinical symptoms of ariboflavinosis.

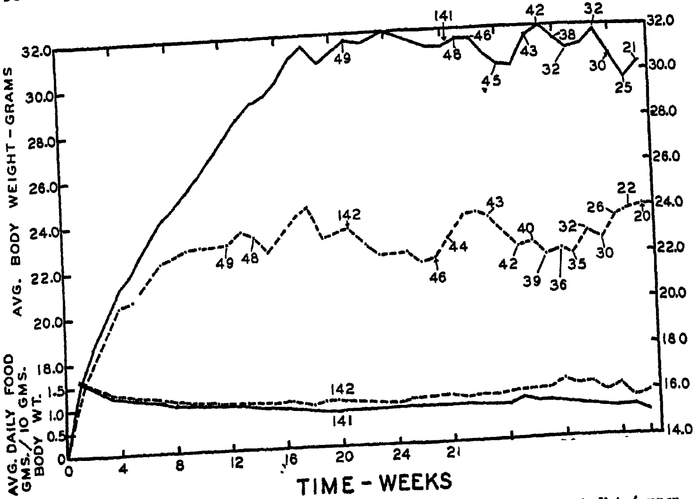


FIGURE 12. Average body weight of mice fed a partially riboflavin-deficient diet (upper curve 142) and of control litter mate mice fed adequate riboflavin supplement (upper curve 141). Two lower curves represent the average daily food intake, respectively, of the two groups. The numbers at various points indicate the number of animals used to obtain the average values.

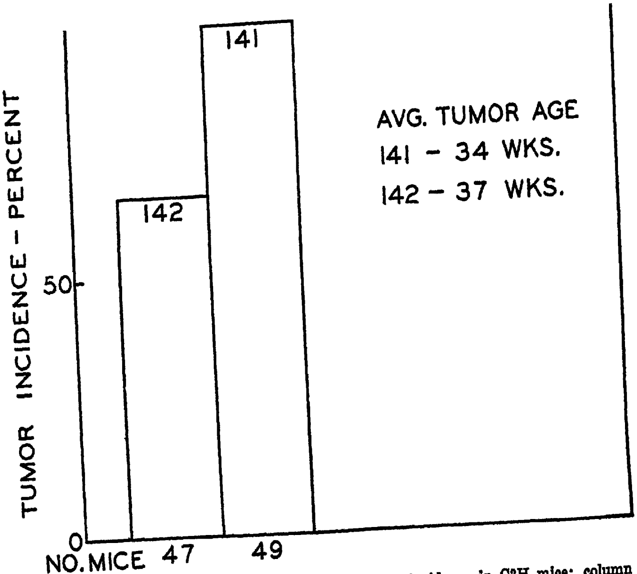


FIGURE 13. Average spontaneous mammary tumor incidence in C3H mice: column 142, for mice ingesting a partially riboflavin-deficient diet; column 141, for mice ingesting a riboflavin-supplemented diet.

The mice partially depleted in riboflavin maintained an average body weight approximately 70 per cent that of the litter mate control animals, as shown by FIGURE 12, body weight curves marked 141 and 142. Both the control and deficient groups were allowed to eat of their food *ad libitum*. The partially deficient mice ate slightly more food per unit of body weight than did the control animals, as shown by the food intake curves 142 and 141, respectively, at the bottom of FIGURE 12.

The incidence of tumors was 66 per cent in the deficient group, compared to an incidence of 98 per cent for the controls (FIGURE 13). This difference by the Chi square test was significant. The average age ot appearance of tumors was 37 weeks for the treated animals and 34 weeks for the controls, a difference not considered significant.

During the fifth month of the experiment, no difference was found in the occurrence of estrus cycles in either deficient or control animals. Yet, the mammary glands of the riboflavin-deficient mice showed a considerable degree of underdevelopment compared to the controls, although there was a considerable amount of variation from animal to animal. This was evidenced by less complex branching than in the controls, with a tendency in glands of deficient animals toward the development of fewer hyperplastic nodules. These structures are generally considered to be the precursors of the mammary tumors. This underdevelopment of the mammary glands, together with the development of fewer hyperplastic nodules, therefore, seems to be an explanation for the decreased incidence of spontaneous mammary tumors in the riboflavin-deficient mice.

DISCUSSION

The rate of growth of the spontaneous mammary tumor was not altered in pyridoxine-deficient mice.[4] The absence of added dietary pantothenic acid in experiments definitely depressed the rate of growth of the spontaneous tumor.[3] On the other hand, Kline *et al.*[8] concluded, from a study of transplanted tumors on low pyridoxine diets in both rats and mice, that "on diets deficient in this vitamin the growth rate of tumors in most animals tends to be diminished." Bischoff *et al.*[9] did not find this vitamin necessary for the growth of transplanted mouse sarcoma 180.

Jones[10] has reported that thiamine deficiency decreased the growth rate of the Jensen rat sarcoma, but the relative importance of thiamine and calories *per se* was not clearly evaluated by him. This effect of caloric restriction was found to be the main explanation for decreased rates of growth of the spontaneous mammary tumor in mice allowed to eat the deficient diet *ad libitum*. It was only by forcibly feeding both thiamine-deficient and thiamine-supplemented diets that a depressing effect of thiamine on mammary tumor growth appeared.

These diverse growth responses of different tumors to some of the vitamins demonstrate the difficulty of making generalizations on the *in vivo* requirements of malignant tissue.

The fact that, in our experiments,[4] tumor-bearing C3H mice on a pyridoxine-deficient diet did not lose weight while non-tumor-bearing adult mice did, may be partially explained by the continued growth of the tumor. Does the tumor obtain pyridoxine from the host? Or does the tumor produce enough pyridoxine for its own needs as well as the pyridoxine requirements of the host? Unfortunately, no quantitative determinations were made of the pyridoxine content of either the host or the tumor. Schweigert *et al.*[11] have shown, however, that the pyridoxine content of the normal mouse decreases rapidly on a pyridoxine-deficient diet.

VITAMIN AND HORMONE RELATIONS. The food intake per unit of weight in the mice maintained from weaning time to death in a state of partial ariboflavinosis,[4] was slightly greater than for the controls. The body weight, however, was approximately 30 per cent less than for control mice. Since our observations indicated no effect on estrus cycles, it would appear that partial depletion of the animal in riboflavin did not greatly interfere with the ovarian secretion of estrogen. Another possibility could be that, although estrogen secretion may have been depressed, prolonged partial riboflavin depletion interfered with the normal inactivation of the endogenous estrogen that was produced, thus giving rise to a high endogenous level. While the partially riboflavin-depleted animals showed little inhibition of estrus, there was noted under-development of the mammary glands. The effect of chronic riboflavin deficiency on the pituitary has not yet been studied. Some of these observed effects may possibly have been mediated through the secretions of that gland.

Estrogenic and pituitary hormones are involved in the formation of spontaneous mammary cancer in the mouse.[7] Several dietary regimens which inhibit the genesis of the mouse mammary tumor appear to exert their effect through the endocrine secretions. Therefore, one extremely important field for future studies on the effect of nutrition on carcinogenesis would seem to be a study of the relation of the vitamins to the elaboration of hormones or secretions of the endocrine glands, and the effect of vitamins in promoting the inactivation of such hormones by the organism.

Thus, the possibility that thiamine deficiency may be implicated in the origin of cancer is suggested by the literature relating to the inactivation of estrogen by the liver, and, in cases of incomplete inactivation, the possibility that high levels of endogenous estrogen may prevail over prolonged periods and influence the development of uterine cancer. It has been established by several investigators[12, 13, 14] that the inactivation of endogenous estrogen is interfered with in certain types of liver damage. This inactivation of estrogen by the liver has been demonstrated by the intrasplenic implantation of estrogen in spayed female rats. Such animals with undamaged livers fail to come into estrus, since the blood leaving the spleen goes to the liver, where the estrogen is inactivated and, thus, does not get out into general circulation. When such animals are made deficient in the vitamin B-complex or are suffering from inanition, they almost immediately come into estrus, showing that the liver is not inactivating the intrasplenically implanted estrogen. Recent experiments of Drill and Pfeiffer[15] suggest that the caloric restriction which accompanies vitamin B-complex deficiency (especially a thiamine deficiency) renders the liver ineffective in inactivating intrasplenically implanted estrogen pellets.

Although all types of liver damage may not interfere with the inactivation of estrogen, it seems worth speculating that the high incidence of uterine cancer in rabbits after administration of estrogen to animals with damaged livers, reported by Green,[16] may have been influenced by a lessened ability of the damaged liver to inactivate estrogen.

Ayer and Bauld[17] have recently observed that, in uterine cancer in the human, a low excretion of administered thiamine was found concomitantly with a high endogenous level of estrogen. If the liver fails to inactivate, or only partially inactivates, the normally produced estrogen, a sufficiently high concentration of estrogen may exist in the body for long periods which could serve as a continual growth stimulant to sensitive tissues of the female genital tract. The supposedly high endogenous estrogen was reported, by Ayer and Bauld,[17] to show a remarkable drop after the administration of B-complex, when the estrogen was determined by the cornification level in vaginal smears. These authors did not report the effect of thiamine or the B-complex administration upon the amount of food eaten by the patients. If these patients were actually deficient in thiamine, it seems likely that they might also have been eating an insufficient quantity of food. Nevertheless, their observations merit further study including (1) a sensitive measure of endogenous estrogen, (2) estimation of thiamine deficiency, and (3) attempts to correlate the two conditions for the purpose of trying to detect a potential cancer-producing condition before the cancer develops.

The underlying mechanisms responsible for the differences in the growth response of the spontaneous mammary tumor to vitamin intake are still quite obscure. The recent experiments of Ryan and Lederberg[18] working with *Neurospora* suggest one possible mode by which changes in an organism may be influenced by the absence of essential nutritional constituents.

The abovementioned authors have isolated, after ultraviolet treatment of *Neurospora crassa,* a mutant strain which is incapable of synthesizing the amino acid, leucine. Occasionally, fragments of the mycelium of the leucineless mutant grow autonomously, irrespective of the available leucine, and may, under certain conditions, overgrow the culture until the carbohydrate of the media is exhausted. These authors have shown, further, that this adaptation depends on reversion of the leucineless gene to an allele capable of mediating the synthesis of leucine.

If we view the cells in the organism as being regulated by their environment, and if growth of the cells be controlled by the nutrients supplied to them—then, should one or more cells become adapted so they could exist independently of one phase of their environment (for example, by becoming able to synthesize some "essential" nutrient which the host still has to obtain from its food supply), autonomous growth of such a changed cell may result and, as Lederberg[19] proposed, offers one explanation for the origin of cancer. If the setting-up of a nutritional regimen deficient in some constituent increases the rate at which such modification processes may occur, they could conceivably become causative factors in the genesis of cancer. It is appropriate to suggest, here, that the genesis of the tumors found in rats after ingestion of choline-deficient diets for several months as reported by Webster,[20] Copeland and Salmon,[21] and by Engel *et al.*[22] could conceivably involve such modification processes.

SUMMARY

Extreme deficiencies of pantothenic acid and riboflavin, produced rapidly during a short period of time, decreased the rate of growth of the spontaneous mammary adenocarcinoma in strain C3H mice. On the other hand, the production of extreme deficiency of pyridoxine under similar experimental conditions did not affect the rate of growth of the mammary tumor.

The rapid depletion of the mouse in thiamine decreased the rate of growth of the spontaneous mammary tumor to the same extent as its food intake was voluntarily restricted. However, in paired tumor-bearing mice forcibly fed equal quantities of thiamine-deficient or thiamine-supplemented food, the average growth rate of the mammary tumors was depressed in the presence of added thiamine.

The feeding of pyridoxine, thiamine, or riboflavin in amounts sufficient to maintain the body weight of the adult non-tumor-bearing animal also furnished a sufficient amount of these vitamins to eliminate any effect on the growth rate of the tumors.

Riboflavin supplementation resulted in increasing the number of tumors which developed. The riboflavin supplementation prior to the appearance of the first tumor, however, was no more effective than giving the supplement after the tumor appeared.

The production of partial riboflavin deficiency by feeding amounts of riboflavin just sufficient to prevent clinical ariboflavinosis depressed the average body weight about 30 per cent, prevented full development of the mammary glands, and reduced the tumor incidence of the partially deficient mice to two-thirds that of the controls.

Some possible explanations of the effects on tumor genesis and growth due to deficiencies of different vitamins have been discussed.

It was pointed out that the possible relation of thiamine intake to inactivation of estrogen in the body needs to be further investigated to see if an estimation of thiamine deficiency can be a useful criterion for detecting a potential uterine cancer-producing condition prior to the appearance of the cancer at that specific site.

BIBLIOGRAPHY

1. **Morris, H. P.**
   1946. Vitamin requirements of the mouse. Proc. Am. Chem. Soc. (Chicago Meeting).
2. **Morris, H. P., & S. W. Lippincott**
   1941. The effect of pantothenic acid on growth and maintenance of life in mice of the C3H strain. J. Nat. Cancer Inst. **2**: 29–37.
3. **Morris, H. P., & S. W. Lippincott**
   1941. Effect of pantothenic acid on growth of the spontaneous mammary carcinoma in female C3H mice. J. Nat. Cancer Inst. **2**: 47–54.
4. **Morris, H. P., & Celia S. Dubnik**
   Unpublished observations.
5. **Dobrovolskaia-Zavakskaia, N.**
   1945. Sur l'effet de l'aneurine (vitamine B^1) sur la croissance des tumeurs spontanés chez les souris. C. R. Soc. Biol. **139**: 494–495.
6. **Morris, H. P., & Wm. B. Robertson**
   1943. Growth rate and number of spontaneous mammary carcinomas and riboflavin concentration of liver, muscle, and tumor of C3H mice as influenced by dietary riboflavin. J. Nat. Cancer Inst. **3**: 479–489.
7. 1945. Mammary Tumors in Mice. AAAS Monograph **22**. Washington, D. C.
8. **Kline, B. E., H. P. Rusch, C. A. Baumann, & P. S. Lavik**
   1943. The effect of pyridoxine on tumor growth. Cancer Res. **3**: 825–829.
9. **Bishop, F., L. P. Ingraham, & J. J. Rupp**
   1943. Influence of vitamin B_6 and pantothenic acid on growth of sarcoma 180. Arch. Path. **35**: 713–716.
10. **Jones, J. L.**
   1942. The effect of vitamin B_1 deprivation on the appearance, growth rate, and course of the Jensen rat sarcoma. Cancer Res. **2**: 697–703.
11. **Schweigert, B. S., H. E. Sanberlick, C. A. Elvehjem, & C. A. Baumann**
   1946. Dietary protein and the vitamin B_6 content of mouse tissue.
12. **Biskind, M. S., & G. R. Biskind**
   1942. Effect of vitamin B-complex deficiency on the inactivation of estrone in liver. Endocrinology **31**: 109–114.
13. **Segaloff, Albert, & Ann Segaloff**
   1944. The role of the vitamins of the B-complex in estrogen metabolism. Endocrinology **34**: 346–350.

14. **Shipley, R. A., & P. Gyorgy**
1944. Effect of dietary hepatic injury on inactivation of estrone. Proc. Soc. Exp. Biol. & Med. **57**: 52–55.

15. **Drill, V. A., & C. A. Pfeiffer**
1946. Effect on vitamin B-complex deficiency, controlled inanition and methionine on inactivation of estrogen by the liver. Endocrinology **38**: 300–307.

16. **Green, H. S.**
1941. Uterine adenomata in the rabbit. III. Susceptibility as a function of constitutional factors. J. Exp. Med. **73**: 273–292.

17. **Ayer, J. E., & W. A. G. Bauld**
1946. Thiamine deficiency and high estrogen findings in uterine cancer and in menorrhagia. Sci. **103**: 441–445.

18. **Ryan, F. J., & Joshua Lederberg**
1946. Reverse-mutation and adaptations in leucineless nemospora. Proc. Nat. Acad. Sci. **32**: 163–173.

19. **Lederberg, J.**
1946. A nutritional concept of cancer. Science **104**: 428.

20. **Webster, G. T.**
1942. Cirrhosis of the liver among rats receiving diets poor in protein and rich in fat. J. Clin. Invest. **21**: 385–392.

21. **Copeland, D. H., & W. D. Salmon**
1945. The occurrence of neoplasms in the liver, lungs, and other tissues of rats as a result of prolonged choline deficiency. Am. J. Path. **22**: 1059–1079.

22. **Engel, R. W., D. H. Copeland, & W. D. Salmon**
1947. Carcinogenic effects associated with diets deficient in choline and allied nutrients. Ann. N. Y. Acad. Sci. **49**(1): 49.

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
VOLUME XLIX, ART. 2. PAGES 141-326
FEBRUARY 10, 1948

CHROMATOGRAPHY*

By

HAROLD G. CASSIDY, NORMAN APPLEZWEIG, STIG CLAESSON, VICTOR R. DEITZ, BEVERIDGE J. MAIR, A. J. P. MARTIN, STANFORD MOORE, ROBERT L. PECK, W. A. SCHROEDER, LEO SHEDLOVSKY, WILLIAM H. STEIN, HENRY C. THOMAS, AND L. ZECHMEISTER

CONTENTS

PAGE
Introduction to the Conference on Chromatography. *By* HAROLD G. CASSIDY 143
History, Scope, and Methods of Chromatography. *By* L. ZECHMEISTER... 145
Chromatography: a Problem in Kinetics. *By* HENRY C. THOMAS........ 161
Frontal Analysis and Displacement Development in Chromatography. *By* STIG CLAESSON 183
Some Experiments in Systematic Quantitative Chromatography. *By* W. A. SCHROEDER 204
Fractionation and Analysis of Hydrocarbons by Adsorption. *By* BEVERIDGE J. MAIR 218
Stereochemistry and Chromatography. *By* L. ZECHMEISTER............ 220
Chromatography in the Streptomycin Problem. *By* ROBERT L. PECK...... 235
Partition Chromatography. *By* A. J. P. MARTIN........................ 249
Partition Chromatography of Amino Acids on Starch. *By* STANFORD MOORE and WILLIAM H. STEIN........ 265
A Review of Fractionation of Mixtures by Foam Formation. *By* LEO SHEDLOVSKY 279
Ion-Exchange Adsorbents as Laboratory Tools. *By* NORMAN APPLEZWEIG 295
The Surface Areas of Some Solid Adsorbents of Possible Use in Chromatography. *By* VICTOR R. DEITZ........ 315
Concluding Remarks. *By* HAROLD G. CASSIDY........................ 324

---

* This series of papers is the result of a Conference on Chromatography held by the Section of Physics and Chemistry of The New York Academy of Sciences, November 29 and 30, 1946.

Publication made possible by grants from the Nathaniel Lord Britton Fund, the Ralph Winfred Tower Memorial Fund, and the General Funds of The New York Academy of Sciences.

COPYRIGHT 1948

BY

THE NEW YORK ACADEMY OF SCIENCES

INTRODUCTION TO THE CONFERENCE ON CHROMATOGRAPHY

By Harold G. Cassidy

Sterling Chemistry Laboratory, Yale University, New Haven, Connecticut

The conference which opens this morning is another in the series sponsored by the Section of Physics and Chemistry of The New York Academy of Sciences. The subject of this conference is *chromatography*, and the papers to be presented will deal with the scope, theory, manipulation, and application of this tool.

Dr. Fruton suggested about a year ago that a conference on this subject be arranged. It was felt that it would be profitable to examine the state of our knowledge regarding chromatography, and especially to make it possible for workers using one method of adsorption analysis to come together with those using other methods, so that all methods might become more widely known. It was hoped that, through such a meeting, the chromatographic tool might become further sharpened and the realm of its proper use redefined.

This present conference developed from these ideas. It emphasizes the very fundamental fact that the improvement of an analytical tool is of major importance to the science, for many an improvement in analytical technique has changed the face of scientific theory. This conference recognizes that the labor of the analyst is as endless as it is rewarding. One thinks, in this connection, of Walt Whitman's lines, "For it is provided in the essence of things, that from any fruition of success, no matter what, shall come forth something to make a greater struggle necessary." No sooner has one problem been solved than the very solution of it raises others of greater subtlety. We have gathered here to examine the new demands which success imposes, as well as to enjoy some of its fruits.

We have met here to discuss chromatography. It is only one of many analytical tools. Like so many of the procedures used by chemists to separate mixtures, it depends upon a partition process. In it, the substances to be separated are allowed to distribute themselves between two or more phases, and the operation is completed by a mechanical separation of the phases. In the case of chromatography, the partition occurs between a fluid phase and an interfacial phase. It is, in short, an adsorption process, and this distinguishes chromatography from other separation processes. In another respect, however, chromatography is similar to other separation methods, such as fractional distillation under reflux, countercurrent extraction, and countercurrent crystallization, in that advantage is taken of the partition in the most efficient manner possible, namely through a countercurrent application. It is partly in this respect that partition chromatography belongs here. It is worth emphasizing, perhaps, on the basis of these similarities, that much that will be said in this publication has implications beyond chromatography.

In the spirit of these Conferences, only relatively few papers have been arranged for, so that there will be ample time for pertinent discussion after each paper. This has meant that many interesting subjects (chiefly specific applications) have had to be passed over. However, we have tried to obtain a balance between theory and method on the one hand and application on the other.

It is now time to turn to those experts who, during the conference, will act as our guides, philosophers, and friends. They will guide us by paths many of which they have themselves broken, and they will discover for us, in the discussion of chromatography, principles which are at the foundation of all analysis and, hence, belong to the very philosophy of science. It is hoped that out of this meeting will come the stimulus to many new researches, hypotheses, and discoveries in and through the use of chromatography.

HISTORY, SCOPE, AND METHODS OF CHROMATOGRAPHY

By L. Zechmeister

California Institute of Technology, Pasadena, California

HISTORY

It is one of the characteristic features of the history of organic chemistry that the direction and rate of progress have depended, to a marked extent, upon the availability and effectiveness of physical methods. In the eighteenth century, when the chemistry of the carbon compounds began to emerge from its infancy, crystallization played an important role in laboratory technique and soon developed into fractional crystallization. Since the middle of the nineteenth century, distillation, an art practised for thousands of years, has gradually developed into fractional distillation of very high efficiency without which modern developments in the field of volatile substances would not have been possible. Likewise, in the present century, we are witnessing the development of the adsorption process into manifold methods of fractional adsorption and elution without which a deeper study of many non-volatile compounds would be unthinkable.

While the history of crystallization or distillation cannot be traced to a unique source, chromatography is based on a single principle and is the invention of one individual: the principle is the passing of a solution through an adsorbing column and the inventor is Michael Tswett.

Tswett possibly was influenced by the work of Goppelsroeder, whose experiments on capillary analysis he cites in one of his first papers. Goppelsroeder first described this subject as early as 1861 and, much later, in his well known monograph (1901). The basic observations were, however, made by Goppelsroeder's teacher, Schoenbein (1861, 1864) who observed selective adsorption of the components of a mixture by the different heights to which they rose when a strip of filter paper was dipped into the solution. Schoenbein made his first observations during his classic studies of ozone, when he impregnated paper with potassium iodide-starch or other ozone reagents.

At the age of 34, Tswett, who was the son of a Russian father and an Italian mother, described the fundamental principle and technique of chromatography in an eight-page article (1906) which reached the Editors of the *Berichte der deutschen botanischen Gesellschaft* on June 21, 1906, a date which we may accept as the official birthday of chromatography. In this remarkable paper he tells the story of his discovery.

It was known that various solvents behave very differently in the manner in which they extract the pigment of carefully dried green leaves. Alcohol or ether removes all of the pigments and, in such an extract, the dark green color of chlorophyll obscures the yellow carotenoids. However, petroleum ether yields a yellow extract which contains chiefly carotenoids and is practically free of chlorophyll, in spite of the fact that crude chlorophyll samples, once prepared, are easily soluble in petroleum ether. Tswett also observed that, if as little as 1 per cent of alcohol is added to the petroleum ether, the entire chlorophyll of the leaves goes into solution. He asks the question: "What is the solubility-promoting effect of the alcohol?"

In order to gain some experimental data, Tswett prepared a petroleum ether-alcoholic total extract into which he placed filter paper strips, whereupon the solvent was evaporated. The dried dark-green strips behaved exactly like the leaf powder: pure petroleum ether gave a yellow extract, but the same solvent containing some alcohol yielded a dark green solution and the paper became colorless. Therefore, Tswett concluded:

"The previously puzzling phenomena which were mentioned at the beginning of this article are, consequently, based on the adsorption of the pigment . . . which is overcome by alcohol, ether, etc. but not by petroleum hydrocarbons."

Being a man of imagination, Tswett, forty years ago, developed an astonishingly clear concept of the fundamental processes upon which modern chromatography is based. He wrote:

"An adsorbent which is saturated with one substance is still able to take up and bind a certain amount of another. However, then also substitutions may take place. For example, the xanthophylls are partially displaced by the chlorophyllins from their adsorption compounds; but not *vice versa*. There exists a certain adsorption sequence, according to which the substances are able to replace one another. The following important application is based on this law. If a petroleum ether solution of chlorophyll is filtered through a column of an adsorbent (I use mainly calcium carbonate which is stamped firmly into a narrow glass tube), then the pigments, according to the adsorption sequence, are resolved from top to bottom into various colored zones, since the stronger adsorbed pigments displace the weaker adsorbed ones and force them farther downwards. This separation becomes practically complete if, after the pigment solution has flowed through, one passes a stream of pure solvent through the adsorbent column. Like light rays in the spectrum, so the different components of a pigment mixture are resolved on the calcium carbonate column according to a law and can be estimated on it qualitatively and also quantitatively. Such a preparation I term a chromatogram and the corresponding method, the chromatographic method.

"It is self-evident that the adsorption phenomena described are not restricted to the chlorophyll pigments, and one must assume that all kinds of colored and colorless chemical compounds are subject to the same laws. I have so far investigated lecithin, alkannin, prodigiosin, sudan, cyanin, solanorubin, and acid derivatives of the chlorophyllins with positive success."

Evidently, Tswett is the true inventor of chromatography in all of its important aspects. However, when the time is opportune for the clear formulation of a new principle, several minds may simultaneously have the information on which to base it. In this country, thousands of miles from Tswett's laboratory and nine years before him, Day (1897), discussing the origin of Pennsylvania earth oil, wrote:

". . . by experimental work it may easily be demonstrated that if we saturate a limestone, such as the Trenton limestone, with the oils characteristic of that rock and exert slight pressure upon it, so that it may flow upward through finely divided clay, it is easy to change it in its color to oils similar in appearance to the Pennsylvania oils, the oil which first filters through being lightest in color and the following oils growing darker."

Later, Day induced Gilpin and his collaborators (1908, 1910, 1913) to carry out detailed experiments, and they showed, evidently without the knowledge of Tswett's papers, that, if crude oil is forced upward through a column of fuller's earth, the following sequence can be noticed from top to bottom: saturated aliphatic hydrocarbons, then aromatics and unsaturated substances, and, finally, nitrogen and sulfur compounds, the amounts of which increased towards the bottom "because of selective adsorption."

These experiments, originating from considerations which were quite different from Tswett's, might well, under favorable conditions, have developed into systematic chromatography.

As is well known, the significance of an outstanding discovery is not always understood immediately. In many cases, further developments begin only after a considerable period of latency. The length of such a period seems to be a measure of the advance of a master mind over his contemporaries. The latency period of chromatography lasted for 25 years and, during that time, only a few and scattered applications were reported. Tswett's detailed monograph appeared in 1910, in Russian, and, in the subsequent few years, Dhéré and Vegezzi (1916) made use of it. In the United States, Palmer (1922) was certainly one of the first investigators who carried out extensive chromatographic studies, and it is evident from his remarkable monograph on carotenoids that he realized the importance of the method. Although Tswett had demonstrated the duality of chlorophyll, it is a historically interesting fact that the great pioneer in that field, Willstätter, so far as is known, never employed chromatography.

The latency period of chromatography was broken in 1931, by Kuhn and Lederer as well as by Kuhn, Winterstein, and Lederer (1931) who, working on a preparative scale, resolved plant carotene into several components, and fulfilled Tswett's prophecy of 21 years before, that

"very likely carotene is not a chemical entity but a mixture of two or more homologues which it may be possible to separate from each other by means of adsorption methods . . ."

Fifteen years have elapsed since the resolution of carotene, and still chromatography is essentially empirical in character. The mathematical analysis of its procedures, as carried out by competent authors, so far has had scarcely any influence on practical experimentation.

One of the factors which, at present, hinder a more rapid theoretical development is the lack of a great number of commercially available, cheap adsorbents, expertly packed, promptly delivered, and of constant and guaranteed quality. By quality, of course, I do not mean a grade or a sieve number on the label, but rather the statement of the result of a careful standardization, based, for example, on the behavior of the material in the presence of well-selected dyes, as suggested by Brockmann and Schodder (1941). In spite of the increasing importance of chromatography in some commercial processes, industrial circles have so far shown very little interest in this problem. The present deplorable state of affairs reminds us of that in Bunsen's laboratory a century ago, when each scientific worker first had to make and calibrate his burets and pipets before he could begin to work.

SCOPE

In spite of many hindrances, more than a thousand papers have appeared, during the last decade, in which successful chromatographic experiments are reported. These many data make it possible to evaluate the scope of the method. For more detailed information, recent monographs, with bibliographies, by the following authors may be consulted: Brockmann (1943); Cook (1941); Hesse (1943); Strain (1945); Vetter (1939); Williams (1946); Willstaedt (1939); Zechmeister and Cholnoky (1943).

Although chromatography seemed, earlier, to be of special advantage to investigators of naturally occurring substances, we are now witnessing its introduction into syntheses. The purification and check of purity of artifacts by chromatographic filtration will, for example, probably soon become a routine operation. Its evident advantage over other procedures is that pigmented or fluorescing by-products and contaminants may easily be located on the column and identified if necessary. Such a process also offers an efficient analytical check for the standardization of industrial products.

Other well-known applications of chromatography can be placed under the headings: resolution of a mixture, accumulation of substances which occur in great dilution, and checking of the homogeneity of a compound as well as its identity or non-identity with another sample by a mixed chromatogram test.

Location of Zones. Since the majority of carbon compounds are colorless, a rapid and reliable method for the location of invisible zones is often required. Under favorable conditions, no special procedures may be necessary for this purpose. For example, Trappe (1941) reports that silicic acid, which appears to be translucent when moistened with a pure solvent, loses this property at the section where a zone is present.

The inspection of the column with a portable ultraviolet lamp during the chromatographic process, either in a dark room or under a black cloth, is becoming increasingly popular (Karrer and Schöpp, 1934; Winterstein and Schön, 1934). On the other hand, strongly fluorescent columns allow the location of numerous compounds which are able to quench the fluorescence of the adsorbent and appear as dark zones (Sease, 1947; Brockmann and Volpers, 1947). Claesson (1947) suggests the use of total reflection on a thick glass plate for the detection of invisible zone boundaries.

Insufficient use has been made, so far, of the brush method, *i.e.*, streaking the column with a brush which has been dipped into a suitable reagent (Zechmeister, Cholnoky, and Ujhelyi, 1936). Of course, such streaks can also be inspected in ultraviolet light. Considering the immense number of well-defined color reactions, there is almost no limit to their chromatographic application. For many unsaturated compounds, permanganate may be used, as has been shown, *e.g.*, for stilbene (Zechmeister and McNeely, 1942), and, especially, for a number of sugar derivatives and free carbohydrates by Wolfrom and his collaborators (Lew, Wolfrom, and Goepp, 1945, 1946; McNeely, Binkley, and Wolfrom, 1945; Binkley and Wolfrom, 1946; Georges, Bower, and Wolfrom, 1946).

The applications of color reactions may be carried out in other ways. Reagent paper strips can be attached to the lower end of the chromatographic tube for this purpose. The appearance of a positive reaction then indicates that a zone is just entering the filtrate. One can also test for the location of some compounds by applying a reagent to samples from different sections of the extruded column.

Determination of the position of zones containing biologically active substances may be made by means of special tests of various descriptions. Haller, Acree, and Potts (1944) prepared extracts of the abdominal tips of the virgin female gypsy moth and, after chromatography, located that section which contains the sex attractant by exposing evaporated eluates of empirically cut fractions to male moths. One fraction attracted 114 males, the other only 26.

Sequence of Zones. After having discussed the location of zones, we must necessarily consider their chromatographic sequence. The sequence of two or more given compounds cannot usually be predicted simply on the basis of the individual adsorption isotherms. As was pointed out, for example, for fatty acids (Cassidy, 1940) and for amino acids (Lottermoser and Edelmann, 1938; Schaaf and Reinhard, 1944), the mutual influence of the components of a mixture may constitute an overwhelming factor in determining their sequence. In some cases, predictions seem to be possible on the basis of the permanent dipoles. Arnold (1939) stated that the sequence of decreasing dipoles of *p*-, *m*-, and *o*-nitrophenol is also the sequence of decreasing adsorbabilities. However, no generalization seems to be possible on this basis.

In each case, the absolute and relative adsorbability of a compound will, of course, be a function of the molecular structure, although not every structural detail will be effective in determining the behavior on the active surface. The location of those groups which are responsible for the adsorption behavior in a given system will be of great importance for future theoretical research. This is a problem similar to that of the determination of haptene groups which are responsible for the fixation of a drug on an adsorbing tissue. Within a complicated molecule, several atomic groups may compete for the surface, and it seems that various orientations of the molecule in different chromatographic systems are possible.

If all compounds which are on the same column possess analogous structures (for example, if they are homologues or close analogues), then a change in the adsorbent and developer probably will affect each component in the same direction, and their relative sequence will remain unaltered in the new system.

However, if different functional groups or other structural features are present in the molecules of the compounds to be chromatographed, then the individual responses to a change of the system may be unequal. The adsorption affinity of one compound may be increased and that of another may be decreased, and these changes may even, under favorable circumstances, cause an inversion of the original sequence on the column.

The factors which influence the chromatographic sequence were discussed by Strain (1946) in a new and important communication (*cf.* also Strain, 1942).

Among many other observations, Strain, Manning, and Hardin (1944) found that, when petroleum ether was used as a developer, the sequence from top to bottom of the two chlorophylls and a carotenoid-alcohol, fucoxanthin, on sugar was:

chlorophyll *b*
chlorophyll *a*
fucoxanthin *a*.

In contrast, the addition of as little as 0.5 per cent propyl alcohol to the petroleum ether produced the following sequence:

fucoxanthin *a*
chlorophyll *b*
chlorophyll *a*.

Evidently, the two chlorophylls responded similarly to the change in the system and, hence, their relative sequence remained unchanged. Fucoxanthin, however, which is structurally very different from the chlorophylls, responded differently and moved to the top position in the new system.

As pointed out by Strain *et al.*, if an inversion can be produced simply by changing the ratio of two solvents, then there must exist a mixture which will cause no separation at all.

Another characteristic example of an inversion in which not the developer (benzene) but the adsorbent was changed was given by LeRosen (1942). It refers to the pair, lycopene, $C_{40}H_{56}$, which contains 13 double bonds (all in an open chain) but no hydroxyl, and cryptoxanthin, $C_{40}H_{56}O$, which contains 11 conjugated double bonds (two in rings) and also a hydroxyl group. On alumina or calcium carbonate, the sequence is:

cryptoxanthin
lycopene,

which shows that the presence of hydroxyl is responsible for a strong fixation, and that this effect overrules the presence of a greater number of double bonds in the hydrocarbon. In contrast, the sequence on calcium hydroxide is:

lycopene
cryptoxanthin.

In aqueous systems, a change in pH may cause inversion in a chromatographic sequence.

Interaction between Column and Substance. If we wish to delineate the scope of chromatography, we should also consider how far the adsorption process itself may cause chemical changes. It is common knowledge that, in thousands of experiments, changes did not occur or, to be more exact, that the starting material, after elution, could be recovered quantitatively. On the other hand, to deny in principle the possibility of conversions on the column would be equivalent to the statement that the surface of a powdered solid is unable to react with a solute.

In fact, some interesting chemical changes have been reported. They belong to two different types.

Chemical alterations of the substance may be caused not by an interaction proper between adsorbent and solute, but by the sensitivity of the compound in the adsorbed state toward oxygen, for example. Thus, some unsaturated fatty acids or amino acids undergo autoxidation on alumina or charcoal. This unwelcome process may be checked by degassing the adsorbent and then excluding air during all operations, or, most simply, in some cases, by a pre-treatment of the column material with an agent which counteracts the tendency toward oxidation.

A treatment with potassium cyanide or hydrogen cyanide will prevent catalytic autoxidations which are caused by traces of some heavy metals

(Tiselius, 1941; *cf.*, for example, Turba, Richter, and Kuchar, 1943; Schramm and Primosigh, 1943). It may also happen that such oxidations are promoted by light. Levy and Campbell (1939) observed, when chromatographing 2,3-benzanthracene, that a zone became orange on the side facing the window because of local photo-oxidation to a quinoid compound.

Concerning the chemical interactions proper between substance and column, the available experimental data can be grouped into two subclasses, although the border lines are far from well-defined. In a number of cases, we do not know exactly what happens in the column, and are unable to give definite chemical equations. It is well known, for example, that chlorophyll suffers an irreversible change on alumina, calcium carbonate, sodium sulfate, fuller's earth, but can be chromatographed satisfactorily on sugar or inulin (Winterstein and Stein, 1933; Mackinney, 1938). Vitamin K_1 is biologically inactivated by magnesia or alumina, but is stable on sugar (Dam and Lewis, 1937).

Of several well-defined reactions which occur on the column, the following may be cited. Dimerization is observed if acetone is filtered through activated alumina, the product being diacetone-alcohol (Hesse, Reicheneder, and Eysenbach, 1938). Polymerization and structural isomerization of a number of essential oils have been described by Carlsohn and Müller (1938). Such a catalytic effect of the frankonite column is increased by dehydration of the adsorbent, *in vacuo,* over phosphorous pentoxide. Possibly, other similar changes in labile structures are due simply to exceptionally high heats of adsorption.

Hydrolytic elimination of acyl groups may occur on alumina when triglycerides are chromatographed (Trappe, 1940). Cahn and Phipers (1937) reported that, on alkaline alumina, a hydrolysis of diacetyltoxicarol takes place, and that the reaction may be prevented by pre-treating the adsorbent with acetic acid or phenol. These treatments, however, diminish the adsorbing capacity.

An elimination of water or alcohol may occur on the column. When rhodin-g_1-trimethylester was developed with methanol on talc, acetal formation took place because of the acidity of the adsorbent. However, on neutral talc, no such conversion occurred (Fischer and Conrad, 1939). According to Brockmann and Junge (1943), alcohol was split off by alumina from some anthocyanidine-like synthetic pigments, a reaction which also takes place when the solution is refluxed instead of being chromatographed.

Column reactions are demonstrated conspicuously when color is produced by the adsorption of a colorless substance, or when the color of the starting material is fundamentally altered. Sometimes, color is evident in the presence of a non-polar developer but disappears on subsequent application of a polar solvent. Triphenylcarbinol is adsorbed by alumina from its colorless benzene or chloroform solution with yellow color and by silica gel with brownish-yellow color (Weitz and Schmidt, 1939).

A well-known reaction, detected by Takahashi and Kawakami (1923), takes place when polyene pigments such as carotenoids, faintly colored polyenes such as diphenyloctatetraene, or colorless polyenes such as vitamin A or phytofluene, are adsorbed from petroleum ether or benzene on acid earths, *e.g.*, filtrols. A dark blue or greenish-blue zone appears immediately at the top of the column, but disappears upon the elution with alcohol or acetone. The formation of color may be irreversible, in the sense that after elution the original extinction curve, for example, of phytofluene no longer is evident (Zechmeister and Sandoval, 1945). The formation of color on acid earths was explained recently by Meunier (1942) as a donation of unshared electrons by the compound to those atoms of the adsorbent which had possessed incomplete octets. The molecule thus suffers far-reaching polarization on the adsorbent and forms strongly resonating structures.

The appearance of an orange pigment on calcium hydroxide or alumina columns during the chromatography of vitamin A has been observed in several laboratories (*cf.* Castle, Gillam, Heilbron, and Thompson, 1934). In this connection, Holmes and Corbet (1939) warned against "the blind use of adsorption columns". According to Meunier and Vinet (1945), the formation of the orange pigment is the result of the elimination of one mole of water from two molecules of the vitamin with the formation of a dipolyene ether.

The possibility of such complications should not discourage the chromatographer, since an unwelcome disturbance in experimentation may soon become a new and useful tool. In fact, some changes which occur on the column can be conveniently used for preparative purposes, as, *e.g.*, for freeing a base from its salt, for conversion of one salt into another, for splitting of addition compounds, etc.

Thus, some curare alkaloids, which were available to Wieland and Pistor (1938) in the form of their anthraquinone sulfonates, were filtered through HCl-pretreated alumina. The anthraquinone sulfonic acid was retained on the column and was replaced by hydrochloric acid in the solution, whereupon the hydrochloride of the alkaloid could be crystallized from the chromatographic filtrate. However, on basic alumina, the free alkaloid was obtained. According to Plattner and Pfau (1937) as well as other investigators, the cleanest method for the liberation of many terpenes, sterols, etc., from their picrates or trinitrobenzoates is a filtration through alumina. The nitro derivative forms a colored zone, while the freed colorless component passes into the filtrate. Similarly, as was reported by Weygand and Birkofer (1939), the prosthetic group of "old" yellow enzyme is retained on frankonite, but the protein component is washed into the filtrate.

As far as I know, no use has yet been made of the following suggestion of Tswett: "For special aims . . . one will purposely use chemically effective adsorbents (hydrolyzers, reducers, oxidizers)."

May I mention that chromatography can also influence the course of organic chemical work in many indirect ways. A rather interesting case was reported by Kuhn and Ströbele (1937), whose crude 2-nitro-4, 5-dimethylaniline sample, which easily formed glucosides, lost this ability upon chromatographic purification, as it turned out, because a specific catalyst was eliminated by the column. This catalyst was ammonium chloride, the addition of which restored the ease of the glucoside synthesis.

METHODS

Although the title of this survey includes also the methods of chromatography, only a short outline can be given. Reference must be made to other contributions to the present Conference.

It seems that, at the present time, there are five methods which promise further important developments:

(1) A characteristic feature of recent experimental work is the further perfection of the liquid chromatogram procedure, *i.e.*, the fractional washing of individual compounds into the filtrate. Especially, in the field of di- and tri-terpenes or steroids, including sex hormones, such methods have been worked out in the laboratories of Reichstein and of Ruzicka with great success. They are reported in a great number of papers which appeared in the *Helvetica Chimica Acta* during recent years. In some cases, more than a dozen solvents or solvent mixtures with increasing developing and eluting power were applied successively and, thus, permitted the use of a relatively small column for a large-scale experiment. The very great solvent requirement is an inconvenience but, on the other hand, the fractionation is satisfactory even if the individual compounds are not present in distinct zones but are washed out one by one from a mixed adsorbate. So far as the author knows, no general survey is available in this field.

(2) A further, promising direction of research is the use of exchangers of the zeolite type as column materials. The practical difference in the function of an adsorbent proper and an exchange adsorbent is that, in the latter instance, the process necessarily involves a release of some constituent from the solid surface to the solvent. In most cases, aqueous or water-containing media have been used, and the process can be characterized as an ion exchange. A new stimulus to exchange chromatography has been given by the study of synthetic resins, the selective exchange properties of which were discovered by Adams and Holmes (1935). These procedures are now being used extensively, for example, in the separation of amino acids of different types. The method has

much in common with ordinary chromatography and, thus, one may discuss the observed phenomena in terms of exchange-affinity, exchange-displacement, exchange-chromatographic sequence, and so forth. After adequate conditioning, either cation or anion exchange can be carried out. I may refer to the review by Myers (1942).

(3) Another important experimental technique, *Partition Chromatography,* was introduced with success by Martin and Synge (1941; *cf.* also Gordon, Martin, and Synge 1943, 1944, and connected papers). In this technique, the solid column is essentially a support for a static liquid phase, while a mobile liquid phase flows through the adsorbent. In the usual form of chromatography, a very great number of adsorptions and elutions is responsible for the resolution of a mixture, but in partition chromatography a very great number of consecutive partitions of the substance occur between the two liquid phases. Thus, differences in individual partition coefficients play a decisive role in the success of the experiment. The authors mentioned worked with silica gel and separated acetamino acids. However, it was shown by Synge (1944) that, on starch, free amino acids can be handled as well. A competent survey of this whole field has been given by Martin and Synge (1945).

(4) In connection with the investigations just mentioned, the use of paper strips for the resolution of mixtures as initiated by Schoenbein and by Goppelsroeder and surveyed by Reinholdt (1925), has assumed new shape and significance. Recently, Consden, Gordon, and Martin (1944) described their "two-dimensional chromatography" which is based on the partition principle. The liquid phases are, for example, water (in the cellulose) and a mobile developer. A droplet containing a very small amount of wool hydrolysate is placed on the corner of a quadrangular piece of paper and developed first with collidine along one side of the paper and then with a dilute phenol solution in a direction perpendicular to the first. The position of the resulting amino acid spots can be demonstrated by spraying with ninhydrin solution and gentle heating.

(5) Finally, mention should be made of the chromatographic boundary method invented by Tiselius (1940). A solution is forced through an adsorbent under conditions which make possible a continuous optical characterization of the emerging liquid by means of a device which records the refractive indices (Claesson, 1944). The individual compounds have characteristic retention volumes and show a sudden breakthrough. This method is acquiring considerable analytical importance in many fields of organic chemistry and biochemistry. It can be carried out as a "frontal analysis" (without developer), as an "elution analysis" (developed with a solvent), or as a "displacement development" (using a strongly adsorbed displacer). These methods have been reviewed by Tiselius (1942) and by Claesson (1946).

When one surveys all the methods and applications of chromatography which are available at the present time, including those for the separation of isotopes, for inorganic analysis, and for technological research, one receives the impression of an extended and manifold picture. This vista contrasts sharply with that which Tswett was able to see during the fourteen years between the publication of his first article on chromatography and his death. He enjoyed neither personal happiness nor proper working conditions, and he received scarcely any acknowledgment during his lifetime. Furthermore, his life was made difficult and, finally, destroyed by the events of the first World War and its consequences. Nevertheless, one has the definite impression, in reading his papers, that he believed with almost fanatical faith in his pioneer ideas.

The present short survey is dedicated to the memory of Michael Tswett.

BIBLIOGRAPHY*

Adams, B. A., & E. L. Holmes
1935. Adsorptive properties of synthetic resins. I. J. Soc. Chem. Ind. (Trans.) **54**: 1.

Arnold, R. T.
1939. Chromatographic adsorption and dipoles. J. Am. Chem. Soc. **61**: 1611.

Binkley, W. W., & M. L. Wolfrom
1946. Chromatographic isolation of cane juice constituents. J. Am. Chem. Soc. **68**: 1720.
1946. Isosucrose synthesis. J. Am. Chem. Soc. **68**: 2171.

Brockmann, H.
1943. New Methods in Preparative Organic Chemistry. Verlag Chemie. Berlin.

Brockmann, H., & H. Junge
1943. Benzopyrylium compounds. I. *o*-Quinoid blue anhydro bases. Ber Deutsch. Chem. Ges. **76**: 1028.

Brockmann, H., & H. Schodder
1941. Aluminum oxides with gradated adsorption power for chromatographic adsorption. Ber. Deutsch. Chem. Ges. **74**: 73.

Brockmann, H., & F. Volpers
1947. On the chromatographic adsorption II. A new procedure for the separation of colorless substances. Ber. Dtsch. Chem. Ges. **80**: 77.

Cahn, R. S., & R. F. Phipers
1937. Reactions caused by activated alumina. Nature **139**: 717.

Carlsohn, H., & G. Müller
1938. On the chemistry of clay. II. The behavior of ethereal oils and of their components on clay and related materials. Ber. Deutsch. Chem. Ges. **71**: 863.

Cassidy, H. G.
1940. Adsorption analysis. III. Relation between adsorption isotherm and position on the adsorption column. Am. Chem. Soc. **62**: 3076.

Castle, D. C., A. E. Gillam, I. M. Heilbron, & H. W. Thompson
1934. Adsorption experiments with vitamin A concentrates. Biochem. J. **28**: 1702.

---

* This is not claimed to be exhaustive.

Claesson, S.
1944. A Self-Registering Apparatus for Adsorption Analysis. *The Svedberg*: 82. Almquist and Wiksells. Uppsala.
1946. Studies on adsorption and adsorption analysis with special reference to homologous series. Ark. Kemi. Min. Geol. **23A** (1).
1947. A new method for the observation of zones of colorless substances on a chromatographic column. Nature **159**: 708.

Consden, R., A. H. Gordon, & A. J. P. Martin
1944. Qualitative analysis of proteins: a partition chromatographic method using paper. Biochem. J. **38**: 224.

Cook, A. H.
1941. Chromatographic Analysis. Inst. of Chem. of Great Britain and Ireland. London.

Dam, H., & L. Lewis
1937. The chemical concentration of vitamin K. Biochem. J. **31**: 17.

Day, D. T.
1897. A suggestion as to the origin of Pennsylvania petroleum. Proc. Am. Philos. Soc. **36**: 112.

Dhéré, Ch., & G. Vegezzi
1916. On the pigmentary composition of hepatochlorophyll. C. R. Acad. Sci. (2): 399.

Fischer, H., & M. Conrad
1939. Partial oxidation of some chlorophyll derivatives. Liebig's Ann. **538**: 143.

Georges, L. W., R. S. Bower, & M. L. Wolfrom
1946. Chromatography of sugars and their derivatives. J. Am. Chem. Soc. **68**: 2169.

Gilpin, J. E., & O. E. Bransky
1910. The diffusion of crude petroleum through Fuller's earth. Am. Chem. J. **44**: 251.

Gilpin, J. E., & M. P. Cram
1908. The fractionation of crude petroleum by capillary diffusion. Am. Chem. J. **40**: 495.

Gilpin, J. E., & P. Schneeberger
1913. Fractionation of California petroleum by diffusion through Fuller's earth. Am. Chem. J. **50**: 59.

Goppelsroeder, F.
1861. On a procedure to detect dyes in their mixtures. Verhandl. Naturforsch. Ges. Basel **III**: 268.
1901. Capillary Analysis Based on Capillarity and Adsorption Phenomena. Birkhäuser. Basel.

Gordon, A. H., A. J. P. Martin, & R. L. M. Synge
1943. Partition chromatography in the study of protein constituents. Biochem. J. **37**: 79.
1944. Technical notes on the partition chromatography of acetaminoacids with silica gel. Biochem. J. **38**: 65.

Haller, H. L., F. Acree, Jr., & S. F. Potts
1944. The nature of the sex attractant of the female gypsy moth. J. Am. Chem. Soc. **66**: 1659.

Hesse, G.
1943. Adsorption Methods in the Chemical Laboratory with Special Consideration of the Chromatographic Adsorption Analysis (Tswett Analysis). De Gruyter. Berlin.

Hesse, G., F. Reicheneder, & H. Eysenbach
1938. The heart poisons in the Calotropis milk. Liebig's Ann. **537**: 67.

Holmes, H. N., & R. E. Corbet
1939. Catalytic effects of porous powders on pure vitamin A. J. Biol. Chem. **127**: 449.

Karrer, P., & K. Schöpp
1934. Filtrations at lower temperatures. Chromatographic analysis of colorless compounds (ultrachromatogram). Helv. Chim. Acta **17**: 693.

Kuhn, R., & E. Lederer
1931. Resolution of carotene into its components. Ber. Deutsch. Chem. Ges. **64**: 1349.

Kuhn, R., & R. Ströble
1937. On *o*-nitraniline glucosides. Ber. Deutsch. Chem. Ges. **70**: 773.

Kuhn, R., A. Winterstein, & E. Lederer
1931. Contribution to the knowledge of xanthophylls. Z. Physiol. Chem. **197**: 141.

LeRosen, A. L.
1942. A method for standardization of chromatographic analysis. J. Am. Chem. Soc. **64**: 1905.

Levy, W. J., & N. Campbell
1939. Studies in qualitative organic analysis. II. J. Chem. Soc.: 1442.

Lew, B. W., M. L. Wolfrom, & R. M. Goepp, Jr.
1945. Chromatography of carbohydrates and some related compounds. J. Am. Chem. Soc. **67**: 1865.
1946. Chromatography of sugars and related polyhydroxy compounds. J. Am. Chem. Soc. **68**: 1449.

Lottermoser, A., & K. Edelmann
1938. The adsorption of aliphatic amines and of protein cleavage products on alumina from aqueous solution. Kolloid Z. **83**: 262.

Mackinney, G.
1938. Some absorption spectra of leaf extracts. Plant Physiol. **13**: 123.

McNeely, W. H., W. W. Binkley, & M. L. Wolfrom
1945. Separation of sugar acetates by chromatography. J. Am. Chem. Soc. **67**: 527.

Martin, A. J. P., & R. L. M. Synge
1941. A new form of chromatogram employing two liquid phases. 1. A theory of chromatography. 2. Application to the micro-determination of the higher monoamino-acids in proteins. Biochem. J. **35**: 1358.
1945. Analytical chemistry of the proteins. Adv. Prot. Chem. **2**: 1.

Meunier, P.
1942. On the action of the montmorillonite argiles on vitamin A, and the phenomena of mesomerism in the class of the carotenoids. C. R. Acad. Franç. **215**: 470.

Meunier, P., & A. Vinet
1945. On the structure of the pigment obtained by chromatography of axerophtol (vitamin A). Bull. Soc. Chim. Biol. **27**: 186.

Myers, R. J.
1942. Synthetic resin ion exchangers. Adv. Colloid Sci. **1**: 317.

Palmer, L. S.
1922. Carotinoids and Related Pigments. The Chemical Catalog Co. New York.

Plattner, P. A., & A. St. Pfau
1937. On volatile vegetable compounds. V. The preparation of the basic substance of the azulene series. Helv. Chim. Acta **20**: 224.

Reinholdt, H.
1925. Capillary and Adsorption Analysis. In: **Houben-Weyl,** *The Methods of Organic Chemistry* **I**: 291. Thieme. Leipzig.

Schaaf, E., & O. Reinhard
1944. Adsorption behavior of amino acids. Ber. Deutsch. Chem. Ges. **76**: 1171.

Schoenbein, Ch. F.
1861. Verhandl. Naturforsch. Ges. Basel **III**: 249.
1864. Verhandl. Naturforsch. Ges. Basel **IV**: (1).

Schramm, G., & J. Primosigh
1943. On the quantitative separation of neutral amino acids by chromatography. Ber. Deutsch. Chem. Ges. **76:** 373.

Sease, G. W.
1947. The use of a fluorescent adsorbent for the chromatography of colorless compounds. J. Am. Chem. Soc. **69:** 2242.

Strain, H. H.
1942. Problems in chromatography and in colloid chemistry illustrated by leaf pigments. J. Phys. Chem. **46:** 1151.
1945. Chromatographic adsorption analysis. Interscience Publishers, Inc. New York.
1946. Conditions affecting the sequence of organic compounds in Tswett adsorption columns. Ind. Eng. Chem. **18:** 605.

Strain, H. H., W. M. Manning, & G. Hardin
1944. Xanthophylls and carotenes of Diatoms, Brown algae, Dinoflagellates, and Sea Anemones. Biol. Bull. **86:** 169.

Synge, R. L. M.
1944. Analysis of a partial hydrolysate of gramicidin by partition chromatography with starch. Biochem. J. **38:** 285.

Takahashi, K., & K. Kawakami
1923. Chemistry of vitamin A. I. Separation of the effective constituent of liver oil and its properties. J. Chem. Soc. Japan **44:** 590.

Tiselius, A.
1940. A new method for adsorption analysis of solutions. Ark. Kemi, Min. Geol. **14B.** No. 22.
1941. Adsorption analysis of amino acids and peptides. Ark. Kemi, Min. Geol. **15B** (6).
1942. A new method of adsorption analysis and some of its applications. Adv. Coll. Sci. **1:** 81.

Trappe, W.
1940. The separation of biological fatty substances from their natural mixtures by the application of adsorption columns. Biochem. Z. **306:** 316.
1941. The separation of biological fatty substances from their natural mixtures by the application of adsorption columns. III. The separation of phosphorus- and nitrogen-free lipoid fractions. Biochem. Z. **307:** 97.

Tswett, M.
1906. Physical-chemical studies on chlorophyll. The adsorptions. Ber. Deutsch. Botan. Ges. **24:** 316.
1910. Chromophylls in the plant and animal world. Tipogr. Warshawskago utsch. Okr. Warsaw.

Turba, F., M. Richter, & F. Kuchar
1943. Quantitative chromatographic separation of the monoamino-monocarboxylic acids. Naturwiss. **31:** 508.

Vetter, H.
1939. The chromatographic adsorption procedure and its application in organic chemistry. Akadem. Verlagsges. Leipzig.

Weitz, E., & F. Schmidt
1939. Appearance of color upon the adsorption on surface-active substances. Analogy between adsorption and complex formation. Ber. Deutsch. Chem. Ges. **72:** 1740.

Weygand, F., & L. Birkofer
1939. Isolation and purification of pure "old" yellow enzyme from yeast and a new method of reversible splitting. Z. Physiol. Chem. **261:** 172.

Wieland, H., & H. J. Pistor
1938. On the curarin from Calebasse curare. II. Liebig's Ann. **536:** 68.

Williams, T. I.
1946. An Introduction to Chromatography. Blackie & Son, Ltd. London.

Willstaedt, H.
1939. The chromatographic analysis and its applications. Hermann. Paris.

Winterstein, A., & K. Schön
1934. Fractionation and preparation in pure state of organic substances based on the principle of the chromatographic adsorption analysis. III. Z. Physiol. Chem. **230:** 139.

Winterstein, A., & G. Stein
1933. Fractionation and preparation in pure state of organic substances based on the principle of the chromatographic adsorption analysis. II. Z. Physiol. Chem. **220:** 263.

Zechmeister, L., & L. Cholnoky
1943. Principles and practice of chromatography. Chapman and Hall, London; J. Wiley and Sons, New York.

Zechmeister, L., L. Cholnoky, & E. Ujhelyi
1936. Contribution to the chromatography of colorless substances. Bull. Soc. Chim. Biol. **18:** 1885.

Zechmeister, L., & W. H. McNeely
1942. Separation of *cis* and *trans* stilbenes by application of the chromatographic brush method. J. Am. Chem. Soc. **64:** 1919.

Zechmeister, L., & A. Sandoval
1945. The coloration given by vitamin A and other polyenes on acid earths. Science **101:** 585.

CHROMATOGRAPHY: A PROBLEM IN KINETICS

By Henry C. Thomas

Sterling Chemistry Laboratory, Yale University, New Haven, Connecticut

The theory of chromatography has been discussed by several writers.[1-9] For the most part, the problem has been treated from the point of view of instantaneous equilibrium between solution and adsorbent. Walter[8] has, however, given a kinetic treatment of the problem for the case of ion exchange between solution and solid. The complete solution of this latter problem has been given by the present writer.[10] The present paper, after a review of pertinent parts of the equilibrium theory, will develop mathematical descriptions of single solute chromatography for the two simplest reasonable mechanisms of adsorption considered as a kinetic process. The results of some experiments in which rate data were obtained are reported, and are discussed from the point of view of the kinetic theory. The multiple solute problem is mentioned.

THEORETICAL PART

The Equations of Conservation. The equations of conservation for flow processes such as that of chromatography are well known, but are restated here in the interest of obtaining a connected account of the theory. We consider that the experiment is carried out in a uniform tube uniformly packed with adsorbent. Let the distance from the input end of the tube be measured in terms of the mass of the adsorbent contained therein, say x (grams). We will consider only cases for constant input rate of solution or solvent and will denote this rate by $\dot{V}$ (milliliters per minute). Let time be denoted by t (minutes); concentration, by c (millimoles per milliliter); "concentration" of adsorbed material, by q (millimoles per gram); and the free space or pore volume of the adsorbent, by m

---

[1] Wilson, J. N. J. Am. Chem. Soc. 62: 1583. 1940.
[2] Martin, A. J. P., & R. L. M. Synge. Biochem. J. 35: 1358. 1941.
[3] DeVault, D. J. Am. Chem. Soc. 65: 532. 1943.
[4] Weiss, J. J. Chem. Soc. 145: 297. 1943.
[5] Offord, A. C., & J. Weiss. Nature 155: 725. 1945.
[6] Glückauf, E. Ibid. 156: 205. 1945.
[7] Walter, J. E. J. Chem. Phys. 13: 229. 1945.
[8] Walter, J. E. Ibid. 13: 332. 1945.
[9] Glückauf, E. Proc. Roy. Soc. London 186: (1004): 35. 1946.
[10] Thomas, H. C. J. Am. Chem. Soc. 66: 1664. 1944.

(milliliters per gram). The conservation condition demands that the solute entering a layer dx of the tube in time δt be accounted for as follows:

$$c\dot{V}\,\delta t = \left(c + \frac{\partial c}{\partial x}\,dx\right)\dot{V}\,\delta t + m\,dx\,\frac{\partial c}{\partial t}\,\delta t + dx\,\frac{\partial q}{\partial t}\,\delta t;$$

whence

$$\frac{\partial c}{\partial x} + \frac{m}{\dot{V}}\frac{\partial c}{\partial t} + \frac{1}{\dot{V}}\frac{\partial q}{\partial t} = 0, \tag{1}$$

or, for constant flow rate,

$$\frac{\partial c}{\partial x} + m\frac{\partial c}{\partial V} + \frac{\partial q}{\partial V} = 0. \tag{1a}$$

In the formulation of the kinetic problem, it will be convenient to use the equation in which $\dot{V}$ appears explicitly; for the equilibrium theory, the equation in V and x is, of course, adequate.

For most purposes, it will be convenient to change to independent variables more natural to the problem, namely:

$$\left.\begin{aligned} x &= x, && \text{the mass of adsorbent upstream}\\ y &= \dot{V}t - mx && \\ &= V - mx, && \text{the volume of solution downstream}\end{aligned}\right\} \tag{2}$$

In these terms, EQUATION 1 becomes

$$\frac{\partial c}{\partial x} + \frac{\partial q}{\partial y} = 0. \tag{3}$$

The integral form of the conservation condition (EQUATION 3) may be written down at once. Insofar as c and q are piecewise differentiable functions of x and y, EQUATION 3 ensures the existence of the line integral, independent of path,

$$F = \int_{x_1,\,y_1}^{x_2,\,y_2} (q\,dx - c\,dy). \tag{4}$$

If the path of integration can be taken from $x = 0$ to some point at and beyond which c and q vanish, then F is the total amount of solute, adsorbed and in solution, in the column.

The Equilibrium Theory for Single Solutes. Most of the previous discussions of the theory of chromatography have started with EQUATIONS 1 or 3 and have introduced the assumption of instantaneous equilibrium between solution and adsorbent. It has, of course, been realized that, at best, this assumption is an approximation. The idea is formulated by the introduction of the adsorption isotherm,

$$q = f(c). \tag{5}$$

We are here measuring f per gram of adsorbent. EQUATION 3 now becomes

$$\frac{\partial c}{\partial x} + f'(c)\,\frac{\partial c}{\partial y} = 0. \tag{6}$$

Wilson[1] has discussed the solutions of this equation given by $c =$ constant. The method of characteristics[11] enables one to obtain other solutions of the equation. The equations of an integral surface of EQUATION 6, described by the parameter, s, passing through a given initial distribution of c, say $c_i(\tau)$, are

$$\left.\begin{aligned} x &= s + x_i(\tau) \\ y &= f'(c)\,s + y_i(\tau) \\ c &= c_i(\tau) \end{aligned}\right\} \tag{7}$$

in which τ is the variable which maps the initial distribution. These results are subject to the condition that the initial ($s = 0$) distribution of c is continuous and continuously differentiable. As has been shown by DeVault[3] and by Weiss,[4] this solution will eventually give rise to a physically absurd situation corresponding to a many-valued concentration at a point in the tube. This situation has been taken as an indication of the formation of a sharp front in the chromatogram. The difficulty is removed by replacing it with a discontinuity in the concentration at such a point as will give conservation of matter in the tube. For the saturation of an initially empty tube with solution of concentration c_0, the discontinuity is thus placed at the position x_D given by

$$\left.\begin{aligned} Vc_0 &= \left\{mc_0 + f(c_0)\right\} x_D \\ yc_0 &= x_D\, q_0. \end{aligned}\right\} \tag{8}$$

In the discussion of the development or elution process from this point of view, it is noted that, according to the differential equation, the rate of movement of a point of given concentration in the chromatogram is given by

$$\left.\begin{aligned} &\left(\frac{dx}{dV}\right)_c = \frac{1}{m + f'(c)} \\ \text{or}\quad &\left(\frac{dx}{dy}\right)_c = \frac{1}{f'(c)} \end{aligned}\right\} \tag{9}$$

Thus, for an isotherm convex upward, *i.e.*, such that $f'(c_1) > f'(c_2)$ for $c_1 < c_2$, the points of lower concentration move more slowly. According-ing to this idea, a diffuse trailing boundary is formed on development of an initially sharp chromatogram. It should be noted that we are here retaining the solution of the differential equation for the description of the trailing boundary but are discarding it for the leading boundary of the same chromatogram. This procedure may be mathematically arbitrary, but it is physically possible: the same assignment of position of EQUATION 8 as was used in the saturation case will give conservation of matter in the case of elution with pure solvent.

---

[11] Courant, R., & D. Hilbert. Methoden der Mathematischen Physik II: 51-55. Julius Springer. Berlin. 1937.

For comparison, later, with experimental and further theoretical results, we shall write the expression for an equilibrium trailing boundary based on the Langmuir isotherm

$$q = \frac{Kac}{1 + Kc}, \tag{10}$$

namely

$$Kax = (Kc + 1)^2 y. \tag{11}$$

This equation has been given implicitly by DeVault[3] and explicitly by Weiss.[4] In EQUATION 10, the constant is taken so that large K means strong adsorption. The total adsorption capacity is denoted by a. For a chromatogram eluted from a tube initially saturated at concentration c_0,[12] EQUATION 11 states that the effluent concentration will vanish at a volume given by

$$y = Kax. \tag{12}$$

For such a chromatogram, the theory predicts that the effluent concentration will have the value c_0 until y attains the value

$$y_1 = \frac{Kax}{(Kc_0 + 1)^2}, \tag{13}$$

when the concentration will abruptly begin to decrease.

The equilibrium theory of chromatography has the advantage of simplicity and is, hence, of considerable use in the qualitative interpretation of experiments. Quantitatively it has been little tested. It cannot, of course, be expected to apply in the many cases where it is easy to show that equilibrium is not attained. We can hope to gain some insight into the detailed structure of a chromatogram if we put into the theory some reasonable mechanism which will give a lag in the attainment of equilibrium.

A Kinetic Theory of Chromatography. The most obvious way to introduce the time factor into the theory is to take explicit account of the finite rate of adsorption. Let us suppose that the rate is determined by chemical effects, *i.e.*, that diffusion is not a rate-determining factor. Perhaps the simplest way in which this idea can be formulated is by the assumption of a reaction law, not unreasonable in itself, which will lead directly to a Langmuir type isotherm at equilibrium. In this connection, it should be noted that there is no unique such mechanism[18]; a Langmuir type isotherm experimentally found is no guarantee of simple kinetics. We will suppose that the rate of desorption is first order with respect to the concentration of adsorbed material, and that the rate of adsorption is proportional to the product of the "concentration of empty holes" on the adsorbent and the concentration in solution of the material being adsorbed. Thus, we write

---

[12] Holmes, H. N., H. Cassidy, R. S. Manly, & E. R. Hartzler. J. Am. Chem. Soc. 57: 1990. 1935.

$$\left(\frac{\partial q}{\partial t}\right)_x = \dot{V}\left(\frac{\partial q}{\partial y}\right)_x = k_1(a-q)c - k_2\, q, \tag{14}$$

in which k_1 (ml. mmol.$^{-1}$ min.$^{-1}$) and k_2 (min.$^{-1}$) are the velocity constants in terms of which we hope to describe the process. The resulting equilibrium relation is, of course,

$$q = \frac{k_1\, ac}{k_2 + k_1\, c} = \frac{Kac}{1 + Kc}, \tag{15}$$

with

$$K = {}^{k_1}\!/_{k_2}.$$

For very weak adsorbers, or for the early stages of any such adsorption process, EQUATION 14 may be simplified to read

$$\dot{V}\left(\frac{\partial q}{\partial y}\right)_x = k_1 ac - k_2 q, \tag{16}$$

giving the linear isotherm

$$q = Kac. \tag{17}$$

EQUATIONS 16 and 17 are written down explicitly because it will be simpler to develop *ab initio* for this case the kinetic theory of the corresponding chromatographic process. In what follows, the mechanism represented by EQUATION 14 will be referred to as *Langmuir kinetics;* the simpler case given by EQUATION 16 will be termed *linear kinetics.* The formulae resulting from linear kinetics are appreciably simpler than those for the more general case. They will be useful in some applications and will serve as a guide in the application of the more complex formulae.

As we have seen, the conservation equation (3) ensures the existence of a function, *F*, such that

$$dF = q\, dx - c\, dy, \tag{18}$$

with, then,

$$\left.\begin{aligned} c &= -\frac{\partial F}{\partial y} \\ q &= \frac{\partial F}{\partial x} \end{aligned}\right\} \tag{19}$$

---

[18] Fowler, R. H., & E. A. Guggenheim. Statistical Thermodynamics: 427. Cambridge University Press. 1939.

The corresponding differential equations for the chromatographic processes may at once be written down. We have for the linear case

$$\frac{\partial^2 F}{\partial x \partial y} + A\frac{\partial F}{\partial x} + B\frac{\partial F}{\partial y} = 0, \tag{20}$$

and for the Langmuir case

$$\frac{\partial^2 F}{\partial x \partial y} + A\frac{\partial F}{\partial x} + B\frac{\partial F}{\partial y} + C\frac{\partial F}{\partial x}\frac{\partial F}{\partial y} = 0, \tag{21}$$

in which we have written for convenience

$$A = \frac{k_2}{\dot{V}}, \quad B = \frac{k_1 a}{\dot{V}}, \quad C = -\frac{k_1}{\dot{V}} \tag{22}$$

Since we will need them in the application of EQUATION 21, we state here the definitions of two additional constants,

$$\alpha = \frac{k_1 c_0 + k_2}{\dot{V}}, \ \beta = \frac{1}{\dot{V}} \frac{k_1 \ k_2 \ a}{k_1 \ c_0 + k_2}, \tag{22a}$$

such that $AB = \alpha\beta$.

By appropriate substitutions, EQUATIONS 20 and 21 may be reduced to a single form. If we put, in EQUATION 20,

$$F = e^{-(Bx + Ay)} \Psi \ (x, y), \tag{23}$$

there results

$$\frac{\partial^2 \Psi}{\partial x \partial y} = AB\,\Psi. \tag{24}$$

If, in EQUATION 21, we put[10]

$$CF = ln \left\{ e^{-(Bx + Ay)} \ \Phi \ (x, y) \right\} \tag{25}$$

we again get

$$\frac{\partial^2 \Phi}{\partial x \partial y} = AB\,\Phi. \tag{26}$$

In the case of the linear equation (20), the expressions (19) give us the following formulae for the calculation of the concentrations:

$$\left.\begin{array}{ll} & c = e^{-(Bx + Ay)} \left\{ A\Psi - \dfrac{\partial \Psi}{\partial y} \right\} \\ \text{and} & q = -\,e^{-(Bx + Ay)} \left\{ B\Psi - \dfrac{\partial \Psi}{\partial x} \right\} \end{array}\right\} \tag{27}$$

The corresponding formulae for the Langmuir kinetics are

$$\left.\begin{aligned} c &= \frac{1}{C}\left\{ A - \frac{\partial \ln \Phi}{\partial y} \right\} \\ \text{and} \qquad q &= -\frac{1}{C}\left\{ B - \frac{\partial \ln \Phi}{\partial x} \right\} \end{aligned}\right\} \qquad (28)$$

The Boundary Value Problems for Initially Empty and Initially Saturated Columns. Solutions of the differential EQUATIONS 24 and 26 have now to be obtained which reduce to specified initial conditions in the chromatographic column. In this paper, solutions will be obtained for initially empty and initially saturated columns. Thus, we obtain formulae applicable to the effluent from the columns. Experimentally, at least from the point of view of a test of the theory, these are the most interesting cases. Only for such effluents have accurate data been obtained. In the future, it may become of interest to examine quantitatively the structure of the bands within the tube. The corresponding mathematical problem, which has not been carried through, promises to be more complex than the one treated here.

The boundary values of the functions Ψ and Φ are specified as follows.

SATURATION OF COLUMN INITIALLY EMPTY.

$$\left.\begin{matrix} x = 0, & y \geq 0, & c = c_0, \\ x \geq 0, & y = 0, & q = 0. \end{matrix}\right\} \qquad (29)$$

Linear kinetics. The corresponding boundary values of Ψ follow from EQUATIONS 27. Using the condition on the concentration in solution at the input of the column, we find

$$\frac{d\,\Psi\,(0,y)}{dy} \quad - \quad A\,\Psi\,(0,y) + c_0\,e^{Ay} = 0,$$

the solution of which may readily be verified to be

$$\Psi\,(0,y) = (1 - c_0\,y)\,e^{Ay}. \qquad (30)$$

From the equation determining q, we have

$$\frac{d\,\Psi\,(x,0)}{dx} - B\,\Psi\,(x,0) = 0,$$

whence

$$\Psi\,(x,0) = e^{Bx}. \qquad (30a)$$

In the deduction of these the expressions Ψ (0,0) has been put equal to unity, so that $c = c_0$ for $x = 0$, $y = 0$.

Even more simply, in the more general case, we find from EQUATIONS 28 and 29,

$$\Phi\,(0, y) = e^{\alpha y} \tag{31}$$

$$\Phi\,(x, 0) = e^{Bx} \tag{31a}$$

The quantity α is defined in EQUATION 22a.

ELUTION OF COLUMN INITIALLY SATURATED. The boundary values of the functions for the case of the elution by pure solvent of a column initially saturated at $c = c_0$ are specified by

$$\left.\begin{array}{lll} x = 0, & y \geq 0, & c = 0 \\ x \geq 0, & y = 0, & q = q_0 \end{array}\right\} \tag{32}$$

The value of q_0 is given by the isotherm (EQUATIONS 15 or 17) for $c = c_0$. The corresponding sets of values of Ψ and Φ are

$$\Psi\,(0, y) = e^{Ay} \tag{33}$$

$$\Psi\,(x, 0) = (1 + q_0\,x)\,e^{Bx} \tag{33a}$$

and

$$\Phi\,(0, y) = e^{Ay} \tag{34}$$

$$\Phi\,(x, 0) = e^{Bx} \tag{34a}$$

Solutions of EQUATIONS 24 and 26 for the above sets of boundary conditions can be obtained by several means. The method of the Laplace transform leads to a contour integral which is most convenient as a starting point for the discussion of the solution. The classical method of Riemann leads directly to the unique[14] results which must be obtained. The simplest method is to take results which have been obtained for a similar problem[10] and transcribe them for the present case. These results may be summarized as follows. The equation

$$\frac{\partial^2 \varphi}{\partial u\, \partial v} = \varphi \tag{35}$$

is satisfied by the exponential e^{u+v}, by the Bessel function

$$I_0\,(2\sqrt{uv}) = \sum_{m=o}^{\infty} \frac{u^m\, v^m}{m!\; m!} \tag{36}$$

and by the definite integral

$$\varphi\,(u, v) = e^u \int_0^u e^{-t}\, I_0\,(2\,\sqrt{v\,t})\,dt = \sum_{0 \leq n < m}^{\infty} \frac{u^m\, v^n}{m!\; n!} \tag{37}$$

---

[14] Courant, R., & D. Hilbert. *Op. cit.* II: 311–317.

Some of the important properties of the integral $\varphi(u,v)$ are repeated here for convenience:

$$\varphi(u,v) + \varphi(v,u) = e^{u+v} - I_0(2\sqrt{uv}) \tag{38}$$

$$\frac{\partial\varphi(u,v)}{\partial u} = \varphi(u,v) + I_0(2\sqrt{uv}) \tag{39}$$

$$\frac{\partial\varphi(u,v)}{\partial v} = \varphi(u,v) - \frac{\partial}{\partial v} I_0(2\sqrt{uv}) \tag{40}$$

$$\left.\begin{aligned} \varphi(0,v) &= 0 \\ \varphi(u,0) &= e^u - 1 \end{aligned}\right\} \tag{41}$$

and

$$\lim_{u\to\infty} e^{-(u+v)}\varphi(u,v) = 1. \tag{42}$$

Solutions of the linear EQUATION 35 can be constructed from linear combinations of $I_0(2\sqrt{uv})$ and various φ's. The character of the solutions can be modified to fit different boundary conditions by applying to them a suitable operator which commutes with $\partial^2/\partial u\partial v$; the resulting functions remain solutions of the differential equation. Of the infinity of such operators, the most important for our purposes are

$$\frac{\partial}{\partial u},\ \frac{\partial}{\partial v},\ u\frac{\partial}{\partial u} - v\frac{\partial}{\partial v}. \tag{43}$$

For example, the boundary values (EQUATION 30) for the saturation of a column under linear kinetics suggest that we try as a solution

$$\begin{aligned} \Psi &= e^{Bx+Ay} - \frac{c_0}{A}\left(y\frac{\partial}{\partial y} - x\frac{\partial}{\partial x}\right)\varphi(Ay,Bx) \\ &= e^{Bx+Ay} - c_0\Big[\,y\left\{\varphi(Ay,Bx) + I_0(2\sqrt{ABxy})\right\} \\ &\quad - \frac{B}{A}\,x\left\{\varphi(Ay,Bx) - \frac{\partial}{\partial Bx} I_0(2\sqrt{ABxy})\right\}\Big] \end{aligned} \tag{44}$$

This expression fulfills the prescribed boundary conditions and is, therefore, the solution desired. The solution for the elution of a saturated tube (linear kinetics) is

$$\Psi = e^{Bx+Ay} + \frac{q_0}{B}\left(x\frac{\partial}{\partial x} - y\frac{\partial}{\partial z}\right)\varphi(Bx,Ay). \tag{45}$$

With Langmuir kinetics, the appropriate solutions are, for the column initially empty,

$$\Phi = I_0(2\sqrt{ABxy}) + \varphi(Bx,Ay) + \varphi(\alpha y,\beta x), \tag{46}$$

and, for the column initially saturated,

$$\Phi = I_0(2\sqrt{ABxy}) + \varphi(Ay,Bx) + \varphi(\beta x,\alpha y). \tag{47}$$

The Concentrations. With the aid of the differentiation formulae (EQUATIONS 39 and 40), the introduction of the solutions of the preceding paragraph into EQUATIONS 27 and 28 is a simple matter. The results are as follows:

LINEAR KINETICS

Saturation

$$\frac{c}{c_0} = e^{-(Bx + Ay)} \left\{ I_0 (2 \sqrt{ABxy}) + \varphi (Ay, Bx) \right\} \tag{48}$$

$$\frac{q}{q_0} = e^{-(Bx + Ay)} \varphi (Ay, Bx). \tag{49}$$

Elution

$$\frac{c}{c_0} = e^{-(Bx + Ay)} \varphi (Bx, Ay) \tag{50}$$

$$\frac{q}{q_0} = e^{-(Bx + Ay)} \left\{ I_0 (2 \sqrt{ABxy}) + \varphi (Bx, Ay) \right\} \tag{51}$$

The functions involved being all positive quantities, it is immediately apparent how the concentration of the solution leads in the saturation case and lags in the elution case, as compared to the concentration of adsorbed material.

LANGMUIR KINETICS

Saturation

$$\frac{c}{c_0} = \frac{I_0 (2 \sqrt{ABxy}) + \varphi (\alpha y, \beta x)}{I_0 (2 \sqrt{ABxy}) + \varphi (\alpha y, \beta x) + \varphi (Bx, Ay)} \tag{52}$$

$$\frac{q}{q_0} = \frac{\varphi (\alpha y, \beta x)}{I_0 (2 \sqrt{ABxy}) + \varphi (\alpha y, \beta x) + \varphi (Bx, Ay)} \tag{53}$$

Elution

$$\frac{c}{c_0} = \frac{\varphi (\beta x, \alpha y)}{I_0 (2 \sqrt{ABxy}) + \varphi (\beta x, \alpha y) + \varphi (Ay, Bx)} \tag{54}$$

$$\frac{q}{q_0} = \frac{I_0 (2 \sqrt{ABxy}) + \varphi (\beta x, \alpha y)}{I_0 (2 \sqrt{ABxy}) + \varphi (\beta x, \alpha y) + \varphi (Ay, Bx)} \tag{55}$$

In passing, we may note that neglecting $k_1 c_0$ as compared to k_2, *i.e.*, making $\alpha = A$, $\beta = B$, reduces EQUATIONS 52-55 to the corresponding formulae for the linear case.

It should be noted that the results for Langmuir kinetics are essentially more complex than those for the linear case. They involve, explicitly, the initial concentration of the solution with which the column is being or was saturated. Thus, our theory predicts that if and only if the kinetics are linear will the behavior of the chromatogram be independent of initial concentration. This independence will, then, always obtain with sufficiently dilute solutions.

It is of some interest to notice that, in the case of radial flow through a disc of adsorbent of uniform thickness, one obtains the conservation equation (3) in an identical form if the variables are defined as follows:

$x = \pi r^2 d$, total mass of adsorbent within a circle of radius r;

$y = \dot{V} t - m\pi r^2 d$, total volume through a circle of radius r.

Here, m has the same meaning as before; d is the mass of adsorbent per unit area of the disc; and $\dot{V}$ is the total radial volume rate of flow. Thus, at sufficiently large distances from a small central input hole, the chromatogram in a disc will behave in a fashion entirely analogous to a chromatogram in a column if x and y are properly interpreted. Near the input hole, it would be necessary to reexamine the solutions of the equation, since the boundary values are not specified for $x = 0$.

Limiting Cases for Slow Flow. In the descriptions of the chromatographic process just developed, the essential characteristic from an experimental point of view is the explicit dependence of the various concentrations on the volume rate of flow through the tube. For a given set of velocity constants, the adsorption process will approach more and more nearly to equilibrium with decrease in flow rate. This fact finds its expression in the behavior of the various functions for small values of $\dot{V}$. A glance at the definitions 22 and 22a shows that we must examine the asymptotic behavior of the functions for large values of the arguments. This investigation is made easy by the use of an asymptotic expansion of the integral $\varphi\ (u,v)$ due to Professor Lars Onsager.[15] This development runs as follows:

$$\begin{aligned}
\varphi(u, v) &= \frac{1}{2}\left\{1 - H\left(\sqrt{v} - \sqrt{u}\right)\right\} e^{u+v} - \frac{r}{1+r} I_0 \\
&\quad + \frac{1}{2}\left(\frac{1-r}{1+r}\right)\Big[g\left\{2I_0 - \left(I_{\frac{1}{2}} + I_{-\frac{1}{2}}\right)\right\} \\
&\quad + g^2\left\{6I_0 - 4\left(I_{\frac{1}{2}} + I_{-\frac{1}{2}}\right) + 2I_1\right\} \\
&\quad + g^3\left\{20I_0 - 15\left(I_{\frac{1}{2}} + I_{-\frac{1}{2}}\right) + 12I_1 - \left(I_{\frac{3}{2}} + I_{-\frac{3}{2}}\right)\right\} \\
&\quad + g^4\left\{70I_0 - 56\left(I_{\frac{1}{2}} + I_{-\frac{1}{2}}\right) + 56I_1 - 8\left(I_{\frac{3}{2}} + I_{-\frac{3}{2}}\right) + 2I_2\right\}\Big] \\
&\quad + \left(1 - r^2\right)R_5, \qquad (56)
\end{aligned}$$

---

[15] Onsager, Lars. To be published.

in which

$$r = (u/v)^{\frac{1}{4}}, \quad g = r \,/\, (1 + r)^{2}$$

$$I_p = I_p\,(2\sqrt{uv}),\; H(Z) = \frac{2}{\sqrt{\pi}} \int_{0}^{Z} e^{-t^2} dt.$$

R_5 is a small remainder term.
For those values of u and v sufficiently large so that the required accuracy may be obtained by the use of the leading term in the asymptotic expansions of the Bessel functions, the expansion of φ reduces to its first two terms. If, in addition, $|\sqrt{v} - \sqrt{u}|$ is sufficiently large, the error integral may be replaced by the first term of its asymptotic expansion, and we have a much simplified formula for the estimation of the behavior of our various expressions for the concentrations. The results for the effluent of a column may be summarized as follows.

LINEAR KINETICS

Saturation

(a) $Ay < Bx$; that is, $Vc_0 < (mc_0 + Kac_0)x$; the column is not completely saturated.

$$\frac{c}{c_0} \simeq \frac{1}{2\sqrt{\pi}}\, e^{-(\sqrt{Bx} - \sqrt{Ay})^2} \frac{\left(\frac{Bx}{Ay}\right)^{\frac{1}{4}}}{\sqrt{Bx} - \sqrt{Ay}} \tag{57}$$

Here, the concentration is determined by the exponential and is, thus, extremely small in the region considered.

(b) $Ay = Bx$; $Vc_0 = (mc_0 + Kac_0)x$; just sufficient solution has been introduced to saturate the column under equilibrium conditions. Here we make use of EQUATION 38 and find

$$\frac{c}{c_0} \simeq \frac{1}{2} + \frac{1}{4\sqrt{\pi}} \frac{\sqrt{\dot{V}}}{(k_1\, k_2\, axy)^{\frac{1}{4}}} \tag{58}$$

Thus, the concentration ratio differs from ½ by an amount which decreases with the square root of the flow rate. The theory predicts, then, that all saturation curves for moderate rates will cross, for a given column, in the neighborhood of $c/c_0 = ½$.

(c) $Ay > Bx$; $Vc_0 > (mc_0 + Kac_0)x$, *i.e.*, the column is working beyond the "break-through" point.

$$\frac{c}{c_0} \simeq 1 - \frac{1}{2\sqrt{\pi}}\, e^{-(\sqrt{Ay} - \sqrt{Bx})^2} \frac{\left(\frac{Bx}{Ay}\right)^{\frac{1}{4}}}{\sqrt{Ay} - \sqrt{Bx}} \tag{59}$$

In this region, the concentration differs by very little from its initial value.

Elution

For this case, we find similar expressions. If $Ay < Bx$, *i.e.*, insufficient solvent completely to elute the column, we again get EQUATION 59. For $Ay = Bx$, we find, analogous to EQUATION 58,

$$\frac{c}{c_0} \simeq \frac{1}{2} - \frac{1}{4\sqrt{\pi}} \frac{\sqrt{\dot{V}}}{(k_1 k_2 axy)^{\frac{1}{4}}} \tag{60}$$

All elution curves cross in the neighborhood of $c/c_0 = 1/2$. At larger volumes, $Ay > Bx$, EQUATION 57 reappears and we have concentrations exponentially close to zero.

Thus, in the limit for slow flow, the theory for linear kinetics reduces to the equilibrium theory for a linear isotherm. It predicts a chromatogram with sharp front and sharp rear boundary.

LANGMUIR KINETICS

The results in the case of the more complete kinetics are similar. The condition of break-through for the column is

$$\left.\begin{array}{l} \alpha y > \beta x \\ Ay < Bx \end{array}\right\} \tag{61}$$

Under these conditions, we may use EQUATION 42 in our estimate of the functions in EQUATION 52 and find for

Saturation

$$\frac{c}{c_0} \simeq \frac{1}{1 + e^{\frac{k_1}{\dot{V}}(q_0 x - c_0 y)}} \tag{62}$$

Thus, for sufficiently small $\dot{V}$, $c \simeq 0$ or $c \simeq c_0$, according as $c_0 y < q_0 x$ or $c_0 y > q_0 x$. The condition, $q_0 x = c_0 y$, *i.e.*, the condition that the column is just saturated in the equilibrium case, gives the position of the sudden exponential rise in concentration. This result is in agreement with the assignment made necessary in the equilibrium theory by the appearance of the multiple valued concentrations in the leading boundary, EQUATION 8.

In the case of the elution of a saturated column, the limiting agreement of the two theories is even more striking. In the case of the equilibrium theory, the equation for the elution curve may easily be put in the form

$$\frac{c}{c_0} = \sqrt{\frac{y_1}{y}} \frac{\sqrt{Kax} - \sqrt{y}}{\sqrt{Kax} - \sqrt{y_1}} \tag{63}$$

in which y_1 is the value of the effluent volume at which the trailing boundary of the chromatogram appears (EQUATION 13). Inspection of the limiting form of EQUATION 54 for the conditions

$$\alpha y > \beta x$$
$$Ay < Bx$$

shows that the corresponding point is given for $\alpha y \rightarrow \beta x$, $y \rightarrow y_1$, which in view of the definitions of α and β determines the same y_1. After some algebraic jugglery, the limiting form of EQUATION 54 may be put into just the form of EQUATION 63. Thus, again, we have complete agreement in the limit between the two theories. It should be pointed out, however, that for finite flow rate the kinetic theory gives a continuous decrease in concentration from c_0 and an exponential approach to zero at the far end of the tail of the chromatogram. In this respect, it more nearly represents the facts than does the equilibrium theory.

The Determination of the Velocity Constants. In the case of linear kinetics, fairly simple formulae may be obtained for the approximate determination of the velocity constants. Using this simplified picture, we cannot determine k_1 alone; it occurs only in the combination k_1a. The position of the mid-point of the saturation curve serves as approximate determination of Ka, for, by EQUATION 58, this point differs but little from that given by $y = Kax$. An estimate of the magnitude of k_2 may be had from the slope of the saturation curve at $c/c_0 = 1/2$. Differentiation of EQUATION 48 with respect to y gives

$$\frac{\partial\,(c/c_0)}{\partial y} = e^{-(Bx + Ay)} \frac{\partial}{\partial y} I_0 \left(2\sqrt{ABxy}\right). \qquad (64)$$

For sufficiently low flow rate, we may use only the first term of the asymptotic expansion of I_0 and find

$$\frac{\partial\,(c/c_0)}{\partial y} = \frac{A}{2\sqrt{\pi}} \frac{e^{-(\sqrt{Bx} - \sqrt{Ay})^2}}{(ABxy)^{\frac{1}{4}}} \left\{\sqrt{\frac{Bx}{Ay}} - \frac{1}{4Ay}\right\} \qquad (65)$$

which for $Bx = Ay$ reduces to

$$\textit{Slope at Mid-Point} \simeq \frac{1}{2\sqrt{\pi}} \cdot \frac{1}{y} \left\{\sqrt{Ay} - \frac{1}{4\sqrt{Ay}}\right\} \qquad (65a)$$

Thus, we have a means of finding A and, hence, k_2.

In the case of Langmuir kinetics, it is apparent, from EQUATION 62, that the center of the saturation curve will give a measure of K. In this case, since an experimental equation for the isotherm enables one to determine both K and a, one can find values for both velocity constants. Only in the limiting case in which EQUATION 62 is applicable can a very simple expression for the slope of the saturation curve be obtained. One finds easily

$$\textit{Slope at Mid-Point} \simeq \frac{1}{4}\frac{k_1 c_0}{\dot{V}} \tag{66}$$

The applicability of EQUATION 66 will probably prove to be limited. It can be valid only for very slow flow; in these cases, the curves will be steep and the slopes difficult to determine with any precision.

The methods of this paper have thus far been applied only in a limited number of cases. With more experience and more experiments, some more generally satisfactory means of evaluating the velocity constants can, no doubt, be developed.

[**Note.** During the preparation of this paper, the contribution on the same subject of L. G. Sillen[16] has come to the author's attention. Sillen has treated the problem in an essentially identical manner. He has found solutions of the differential equations for several special cases. It is of interest to notice that his solutions turn up as particular cases of the general solutions of this paper. Thus, Sillen's Equations 13 are identical with our EQUATIONS 52 and 53 if, in our equations, we put $k_2 = 0$. Similarly, the solution of Sillen by his "method of final functions" for the general case of Langmuir kinetics, his Equation 35, is identical with our limiting EQUATION 62.

The complete solution of the ion-exchange problem considered by Sillen and Ekedahl[17] has been given.[10]]

EXPERIMENTAL PART

To gain information as to the applicability of the kinetics at the basis of the formulae established under the heading of "Theory", experiments have been carried out on the adsorption of anthracene from cyclohexane solution flowing over activated alumina. Provision was made for varying the rate of flow. Analysis of effluent solutions was carried out spectrophotometrically. It should be emphasized at once that these experiments are of a preliminary nature. Although great care was taken to obtain results of as good accuracy as was possible to the experiment, the design of this experiment was deliberately simplified. Quantitatively, the results cannot be said to be of fundamental significance and are not intended to be so taken. The chief simplification of the experiment was the manner in which the adsorbent was handled. The material was merely taken from a large bottle of commercial activated alumina and introduced directly into the column. In a careful experiment, all precautions should be taken to secure reproducible activity of the adsorbent. Furthermore, in the present experiments, from the nature of the apparatus used, there may be some uncertainty in the total volume of solution passed, due to uncertainty in the volume of dead spaces at both the input and exit ends of the tube. Thus, the absolute positions of the mid-points of the various curves are subject to some error. As has been shown, the positions of these points are of importance in the discussion of the results.

---

[16] Sillen, L. G. Arkiv. Kem. Min. och Geol. 22A (15): 1. 1946.
[17] Sillen, L. G., & E. Ekedahl. Arkiv. Kem. Min. och Geol. 22A (16): 1. 1946.

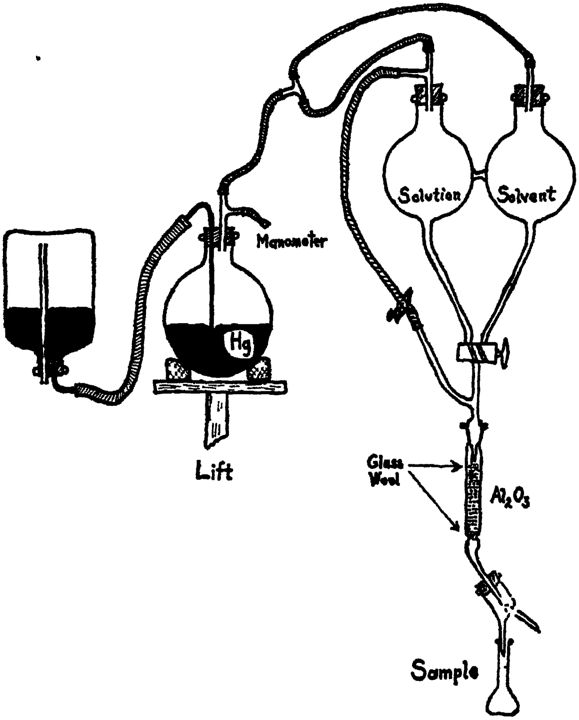


FIGURE 1.

Method and Results. The sketch of FIGURE 1 shows the apparatus used. The pressure was maintained by manual adjustment of the mercury levels in the siphon. The selector stopcock was very carefully ground; no grease was used. The column was filled by sprinkling the alumina into the tube filled with cyclohexane. The alumina was supported on a light wad of pyrex glass wool which had been carefully shaped to give the column as flat a base as possible. The columns were conditioned by forcing a hundred or more milliliters of solvent through them before the experiment was started. Only after such treatment was the rate of flow determined by the applied pressure. During an experiment, the flow rate was maintained constant to better than five per cent.

The entire effluent of the column during an experiment was taken as successive 5, 10, or 25 ml. samples. All solution preparation and sampling were done by weight. Volumes and concentrations were calculated using 0.778 for the density of the solutions.

The experiments were carried out in a temperature-regulated room, no further attempt at temperature control being made. The room temperature varied no more than a degree from 22°C.

A very pure sample of anthracene was used. The material had been codistilled two or three times with ethylene glycol and washed thoroughly with water. It showed the blue fluorescence characteristic of the pure compound. Commercial cyclohexane was treated with fuming sulfuric acid, washed, dried, distilled, and stored over sodium. For the preparation of these materials I am much indebted to Professor Harold G. Cassidy of this Laboratory. The alumina, as mentioned already, was taken directly from a bottle of 80-200 mesh material.

A Beckman spectrophotometer set at a wavelength of 367 mμ with a slit width of 0.079 mm. was used in the analysis of the cyclohexane solution of anthracene. The method is very reliable in the range of 0-0.01 weight percent. As is shown in FIGURE 2, Beer's law is strictly followed in this region. Twenty solutions, prepared from several different stock solutions, varying from 0.000907% to 0.00999%, were used in the calibration. In terms of concentrations (mmol./ml.) the results are well represented by the straight line

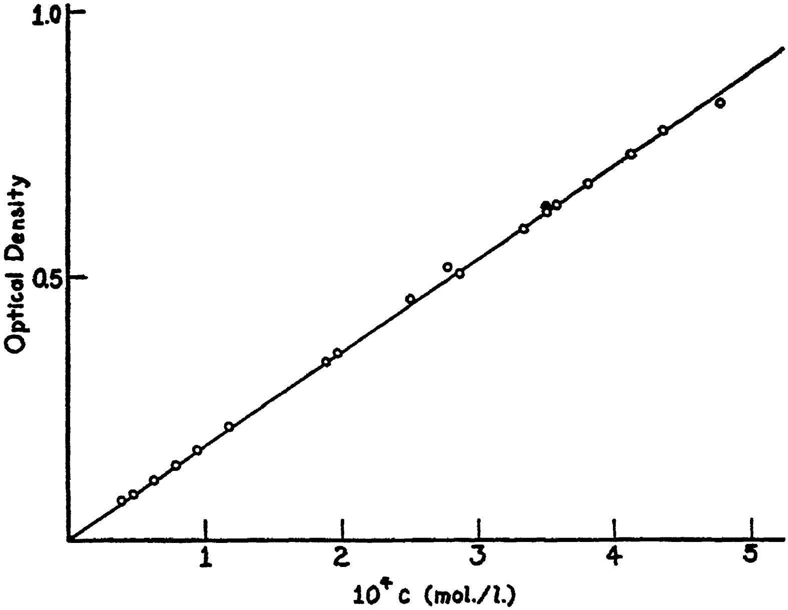


FIGURE 2. Spectrophotometer calibration.

$$c = 5.63 \times 10^{-4} d,$$

in which d is the "optical density" as read directly from the instrument. Except for the most dilute solutions, it is considered that the individual analyses are accurate to 2 per cent or better. Since the concentrations of the input solutions were determined by weight, the ratios c/c_0 are subject to analytical errors no greater than this value.

A determination of the free space of the alumina packed in cyclohexane gave $m = 1.37$ ml./g. Since volumes *through* the column were measured, m does not enter into any of the theoretical computations of the results. The diameter and lengths of the columns are in this same category of unnecessary information. The diameter of the column was 0.82 cm. In the two experiments here reported, the average lengths were 4.79 ± 0.09 cm. and 4.86 ± 0.04 cm., respectively. The weight of alumina in the column was determined by weighing the entire column after the experiment, the solvent having been removed by a current of air.

The results obtained for two complete liquid chromatograms are given in TABLES 1 and 2. These data were obtained using solvent from a single bottle. Unfortunate experience showed that anomalous results might be obtained by using material from slightly different sources. In one elution experiment, the column was saturated using a solution prepared with

TABLE 1
LIQUID CHROMATOGRAM
Anthracene-cyclohexane-alumina at 22° C
Weight of Al_2O_3, $x = 2.316$ g.
Average rate of flow: Saturation = 5.0 ml./min.
Elution = 5.4 ml./min.
Initial concentration, c_0 = 0.000811 mmol./ml.

Saturation		Elution	
Vol. (Δy) ml.	c/c_0	Vol. (Δy) ml.	c/c_0
9.96	0.002	10.06	1.009
9.99	.003	5.04	.990
10.16	.006	4.94	.924
5.06	.038	4.95	.839
5.09	.121	5.12	.734
5.12	.245	4.56	.643
4.91	.403	5.12	.558
9.99	.625	10.05	.431
10.14	.826	10.04	.319
9.97	.925	10.04	.237
10.00	.971	10.07	.183
10.01	.987	9.96	.141
10.07	1.000	10.01	.112
10.01	.998	10.04	.087
10.07	1.006	25.01	.062
9.88	1.003	25.08	.038
		25.01	.022
		25.03	.017
		25.05	.012
		25.12	.007

cyclohexane from one container and eluted with supposedly identical material from another bottle. The initial effluent came through at a much higher concentration than that of the saturating solution, indicating a large change in the character of the adsorption. As a matter of fact, this effect made it necessary to discard the results of several weeks' work. It was impossible to show that the solvent was essentially identical for all the experiments. This sensitivity to very slight impurities in the solvent is another reason why the present results must not be taken as absolute determinations.

The Isotherm. FIGURE 3 contains the results of experiments designed to determine the adsorption isotherm of the system. To weighed flasks equipped with carefully ground stoppers were added, in succession, alumina and appropriate amounts of cyclohexane and stock solution of anthracene in cyclohexane. (All solvent was from the same sample used in determining the chromatograms reported.) After fifteen minutes, during which time the flasks were vigorously shaken at frequent intervals, samples were withdrawn through plugs of cotton wool. Experiments showed that fifteen minutes was ample time for complete reaction. The

TABLE 2
LIQUID CHROMATOGRAM
Anthracene-cyclohexane-alumina at 22° C
Weight of Al_2O_3, x = 2.316 g.
Average rate of flow: Saturation = 9.8 ml./min.
Elution = 9.9 ml./min.
Initial concentration, c_0 = 0.000811 mmol./ml.

Saturation		Elution	
Vol. (Δy) ml.	c/c_0	Vol. (Δy) ml.	c/c_0
10.09	0.000	4.97	1.003
5.10	.003	5.04	.995
5.04	.004	5.02	.941
5.06	.025	5.05	.861
5.07	.071	5.04	.771
10.04	.184	4.93	.678
10.02	.383	10.04	.561
10.04	.582	10.05	.432
10.11	.737	10.04	.332
10.02	.839	10.02	.264
10.07	.905	10.00	.209
9.98	.948	10.02	.167
10.04	.977	9.94	.138
9.97	.983	10.05	.112
10.06	.994	10.03	.093
23.31	1.002	25.11	.070
		24.95	.047
		25.03	.032
		25.08	.023
		25.06	.017
		25.00	.012
		25.00	.010

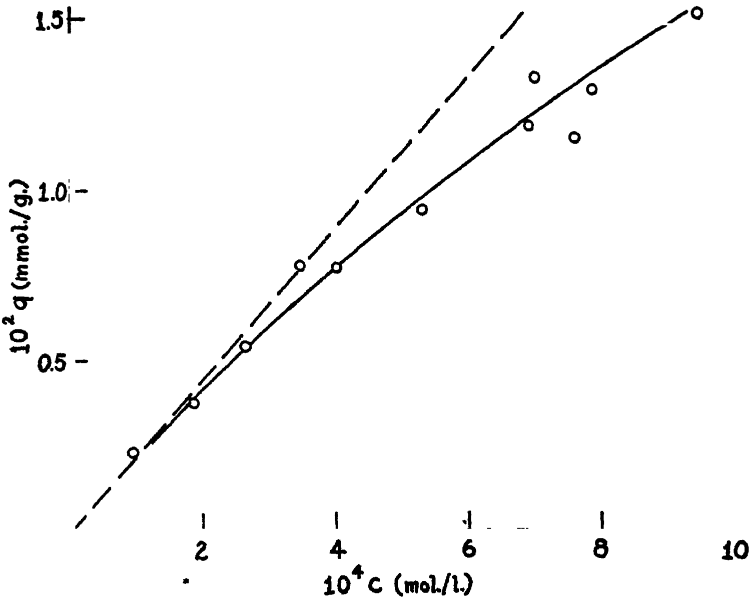


FIGURE 3. Adsorption isotherm $C_{14}H_{10}$ in cyclohexane on Al_2O_3.

cotton plugs in the pipettes were shown to be without effect on the concentration of the samples. Analysis was made in the spectrophotometer. As is apparent, the accumulation of errors involved in determinations of this kind produces large effects on the reproducibility of the results. The solid line of FIGURE 3 is the plot of the Langmuir isotherm,

$$q = \frac{22\, c}{1 + 375\, c}$$

while the dashed line is the plot of the linear isotherm,

$$q = 22\, c.$$

Discussion. It has not been thought worth while to attempt the application of the complete theory for Langmuir kinetics to the experimental results here given. We will content ourselves, at present, with a tentative discussion using the linear kinetics, largely because of the greater simplicity of the computations in this case. After considerable preliminary computation, it has been found that on this basis the results of TABLES 1 and 2 can, perhaps, best be represented if we choose for the constants of the theory,

$$k_1 a = 59.5,$$

$$k_2 = 2.71.$$

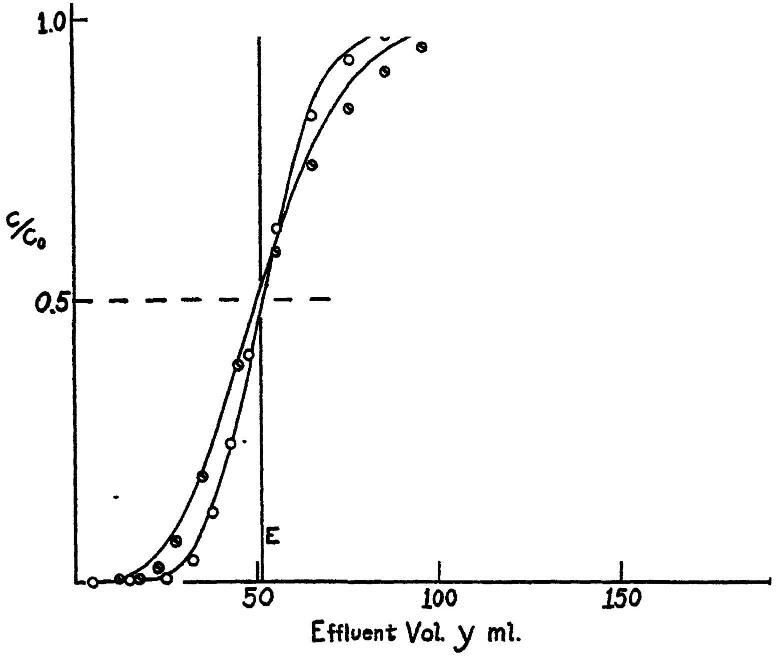


FIGURE 4. Saturation curves $C_{14}H_{10}$ in cyclohexane on Al_2O_3. $O=5.0$ ml./min., $\theta=9.8$ ml./min. Lines from linear theory, $k_1a=59.5$, $k_2=2.71$; E=equil. theory; $q=22c$.

This choice is consistent with the linear isotherm, $q = 22\ c$. In FIGURES 4 and 5, the computed results are given in the form of curves. The experimental points here are plotted against the average y for the particular sample. In the case of the saturation curves, the agreement with the data is better than qualitative. The reproduction of the general character of the results is striking. The quantitative disagreement is, however, in most cases (particularly in the more concentrated solutions toward the upper portions of the curves) certainly outside the limits of a reasonable estimate of the experimental error. In the case of the elution curves of FIGURE 5, we have only qualitative agreement. The theoretical curves lie in the correct order, have the proper general shape, but in no way quantitatively reproduce the data. (They do, however, represent an improvement over the curve given by the equilibrium theory, which is also depicted in FIGURE 5.) In what part these discrepancies are due to fundamental inadequacies of the theory and in what part to overly simplified computation procedure, cannot at present be decided. The first requirement toward making this decision will be more and better experimentation. More experiments are planned for the near future. It is hoped that the theory will also receive testing at the hands of others.

Note on the Theory for Multiple Solutes. The simplest way in

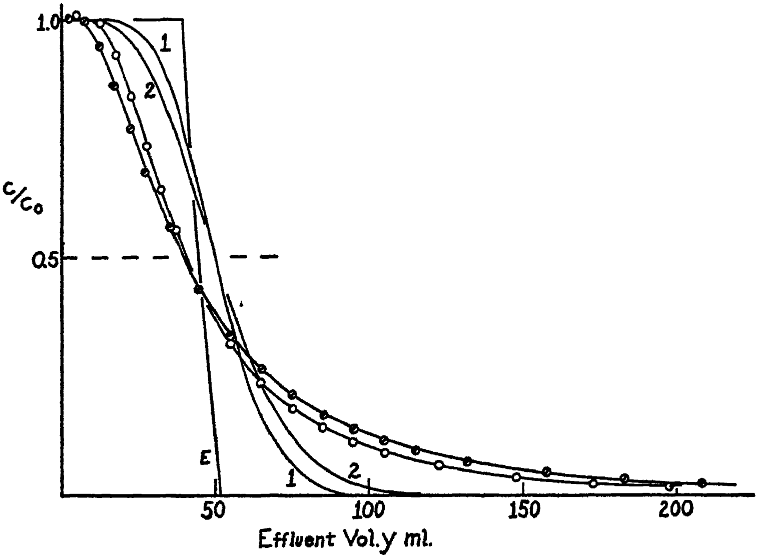


FIGURE 5. Elution curves, $C_{14}H_{10}$ in cyclohexane on Al_2O_3. O=5.4 ml./min., θ=9.9 ml./min. 1 and 2 (same rates) by linear theory, $k_1a=59.5$, $k_2=2.71$; E equil. theory, $q=22c/(1+375c)$.

which interaction during adsorption of two or more solutes can be represented is by neglecting all mutual chemical effects and supposing that the availability of sites on the adsorbent is the determining factor. Again, we predicate the simplest kinetics leading to a Langmuir adsorption isotherm. The differential equations for the corresponding rate processes in the two solute case are,

$$\frac{\partial^2 F_1}{\partial x\,\partial y} + A_1\frac{\partial F_1}{\partial x} + B_1\frac{\partial F_1}{\partial y} + C_1\left\{\frac{\partial F_1}{\partial x}\frac{\partial F_1}{\partial y} + \frac{\partial F_2}{\partial x}\frac{\partial F_1}{\partial y}\right\} = 0,$$

$$\frac{\partial^2 F_2}{\partial x\,\partial y} + A_2\frac{\partial F_2}{\partial x} + B_2\frac{\partial F_2}{\partial y} + C_2\left\{\frac{\partial F_2}{\partial x}\frac{\partial F_2}{\partial y} + \frac{\partial F_1}{\partial x}\frac{\partial F_2}{\partial y}\right\} = 0.$$

The mathematical problem thus posed is most formidable. There seems to be reasonable doubt that the complete solution of such a non-linear system can be obtained by the methods of contemporary mathematics. Possibly, some trick might be found which would reduce the equations to tractable form. After more knowledge of single solute cases is gathered, something might be done with the two-solute case by numerical methods. Presumably, the equilibrium theory of Glückauf[9] forms the limiting cases for slow flow. It is to this theory and to experiment that one must turn for information on what is, after all, the most important case of chromatography, the separation of two substances.

FRONTAL ANALYSIS AND DISPLACEMENT DEVELOPMENT IN CHROMATOGRAPHY

By Stig Claesson

Institute of Physical Chemistry, University of Uppsala, Sweden

INTRODUCTION

In this paper, a description will be given of the arrangements for chromatographic adsorption analysis which have been worked out in Uppsala by Tiselius and his co-workers[1-20].* The basic principle behind the method is liquid chromatography, which means that the whole process of separation is followed by measuring the concentration of the solution leaving the column and not, as in ordinary chromatography, by looking at the zones on the column. The experimental arrangement is given schematically in figure 1. The solution is poured into the vessel *A*, forced through the filter with adsorbent *B*, and the concentration of the solution leaving the filter is observed in the small cell *C*. It is essential that the cuvette *C* should have a very small volume, so as to avoid lag in the readings and to provide for the swift replacement of solution in the cuvette, thus eliminating possible convections. The concentration in the

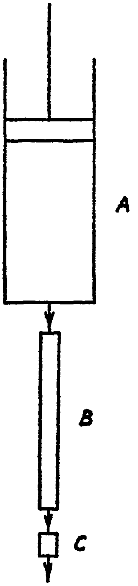


Figure 1. Diagram of apparatus for adsorption analysis.

---

* A detailed description of the whole subject can be found in Claesson, S. Studies on adsorption and adsorption analysis with special reference to homologous series. Arkiv Kemi, Mineral., Geol. 23A, 1: 1–133. 1946.

cell *C* is then plotted against the volume of solution that has passed through the filter *B*. In this way, characteristic curves are obtained from which the qualitative and quantitative composition of a mixture can be calculated.

This procedure has several advantages over the older chromatographic method:

(1) It is especially suited for analysis of colorless substances, and is independent of the color of the adsorbent.

(2) The separation of the different components is much greater in the solution leaving the filter than in the column, which means that the selectivity increases considerably.

(3) The procedure is especially suited for quantitative evaluations of the adsorption phenomenon in the column. It is also clear that, by a proper choice of the method for the determination of the concentration in the cell *C*, it is possible to adapt this method of adsorption analysis for almost all systems met with in practice. In all experiments reported in this paper, the concentration is followed by measurement of the change in the refractive index ($\Delta\mu$) for solutions and the change in the thermal conductivity ($\Delta\lambda$) for gases.

THE INTERPRETATION OF THE DIAGRAMS

When the method for chromatographic adsorption analysis schematically given in FIGURE 1 is used, the experiments can be arranged in three different ways, namely, frontal analysis, elution analysis, and displacement development. A short review of these three adaptations will be given here. For a more detailed description, the reader is referred to Claesson.[18]

Frontal Analysis. Frontal analysis is the simplest method of adsorption analysis and is carried out as follows. The filter with adsorbent is at first washed with pure solvent, after which the solution to be analyzed is poured into the vessel *A* (FIGURE 1) and is allowed to pass the filter *B*. The solute is adsorbed and moves forward with a sharp front as the filter becomes saturated. At the beginning, pure solvent comes through the cell *C*, and the concentration of the solute is zero. When the front has passed the whole filter (that is, when the adsorbent is completely saturated), the concentration in the cell *C* suddenly increases. Diagrams with one step (FIGURE 2) are thus obtained when there is only one solute, with

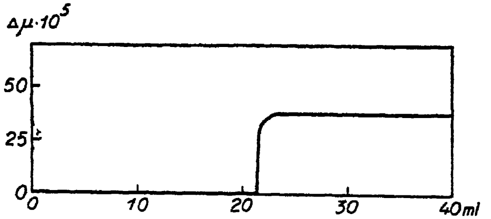


FIGURE 2. Frontal analysis of 0.5% stearic acid in ethanol. Filter: 1250π carboraffin C.

two steps when there are two solutes (FIGURE 3) and so on. The first step contains component 1, the second step component 1 *and* component 2, and similarly the *i*th step components 1, 2, 3, . . . *i*-1 and *i*. The volume that has passed the filter before the solute *i* breaks through was called the *retention volume* for the component *i*, by Tiselius. This volume includes, of course, the small volume of solution necessary to replace the solvent between the adsorbent particles in the filter at the beginning of the experiment. After subtracting this small volume from the retention volume, the *corrected retention volume* v_i is obtained. If the corrected retention volume is divided by the weight of the adsorbent in the filter, the *specific retention volume* v_i^0 (corrected retention volume per gram adsorbent) is obtained.

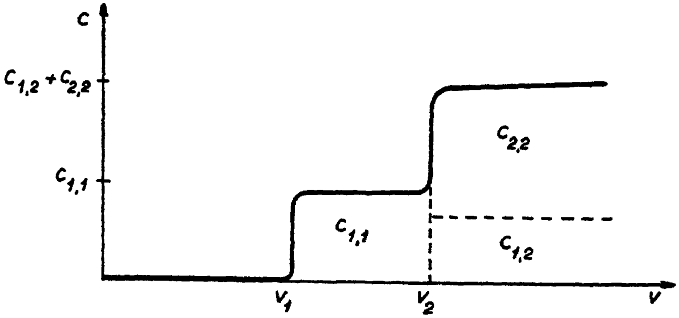


FIGURE 3. Frontal analysis diagram for two solutes.

For a system of one solute, the amount of substance *a* adsorbed in the filter has previously been dissolved in the corrected retention volume *v*, since, after the breaking through of the front, the solution passes unchanged through the filter. If *c* is the concentration of the solution, we get

$$a = v.c \tag{1}$$

or, per gram adsorbent,

$$a^0 = v^0.c \tag{1a}$$

When the adsorbed amount a^0 is plotted against the concentration *c*, the adsorption isotherm *f*(*c*) for the substance in question is obtained, as the amount adsorbed a^0 is in equilibrium with the solution of concentration *c*. Thus, we have

$$a^0 = f(c) = v^0.c \tag{2a}$$

As the adsorption isotherm normally has the shape given in FIGURE 5, the specific retention volume

$$v^0 = \frac{f(c)}{c} \tag{2a}$$

will decrease as the concentration *c* increases. When the concentration tends to zero, the specific retention volume will assume a constant value equal to the slope of the first linear part of the isotherm.

When frontal analysis is carried out on a solution containing two solutes, two steps will appear in the diagram (FIGURE 3). The solution between the two steps contains only solute 1, but its concentration $c_{1,1}$ is higher than in the original solution $c_{1,2}$. This is due to the fact that, as the second front (component 2) moves through the filter, a certain amount of component 1 is displaced from the adsorbent, and appears in the first step.

It is easily seen that, if an experiment is carried out with a solution of known concentrations $c_{1,2}$ and $c_{2,2}$, we have the following expressions for the amounts adsorbed.

$$\begin{aligned} a_1 &= v_2 \cdot c_{1,2} - (v_2 - v_1) \cdot c_{1,1} \\ a_2 &= v_2 \cdot c_{2,2} \end{aligned}$$

or, per gram adsorbent,

$$\begin{aligned} a_1^0 &= v_2^0 \cdot c_{1,2} - (v_2^0 - v_1^0) \cdot c_{1,1} \\ a_2^0 &= v_2^0 \cdot c_{2,2} \end{aligned} \qquad (3)$$

In this way, it is very easy to study mixed adsorption of systems containing two solutes. In the case of three or more solutes, it is not possible to calculate the amounts adsorbed from the frontal analysis diagram, since there are more unknowns than equations.

However, the important thing, from the analytical point of view, is to be able to calculate the concentrations of an unknown solution from the diagram, and that can only be done if the equations of the adsorption isotherms are known. As it is essential to avoid too complicated calculations, a simple approximate formula for the isotherm must be used. It is always possible to carry out all the experiments at the same total concentration, and so even a rough approximation of the isotherm will give rather good accuracy. By using Langmuir's equation for the isotherm,

$$a_i^0 = \frac{k_i c_i}{1 + l_i c_i} \qquad (4)$$

especially simple formulae can be derived. (Here, a_i^0 is the amount adsorbed per gram adsorbent of component i with the concentration c_i; k_i and l_i are constants).

In case of two solutes, the formula will be (FIGURE 2):

$$c_{1,2} = c_{1,1} \cdot \frac{1 - \dfrac{v_1}{v_2}}{1 - \dfrac{k_1}{k_2}} \qquad (5)$$

The correct value for the concentration of component 1 ($c_{1,2}$) is thus obtained directly, as $c_{1,1}$, v_1, and v_2 are easily measured. The concentration, $c_{2,2}$ of component 2 is then obtained by subtraction from the total concentration $c_{1,2} + c_{2,2}$.

In case of n solutes, EQUATION 5 will take the more general form

$$c_{i,m+1} = c_{i,m} \frac{1 - \dfrac{k_i}{k_m} \cdot \dfrac{v_m}{v_{m+1}}}{1 - \dfrac{k_i}{k_{m+1}}} \tag{6}$$

where $c_{i,m}$ is the concentration of component i in the mth step. Consequently, all the concentrations in the diagram can be calculated successively. For the derivation of these formulae, the reader is referred to Claesson.[18]

The expressions obtained are strikingly simple,* and it is remarkable that the constants l_i do not appear, but only k_i, which in homologous series, for example, can be determined by Traube's rule. There, we can put

$$k_n = p.q^n, \tag{7}$$

where p and q are constants, and n represents the number of carbon atoms in the compound. In this way, the calculations are very much simplified, and only two constants p and q need be known for a whole homologous series. Furthermore, a rough estimation of the value k_n can be obtained from the diagram by means of a formula (Claesson[18])

$$k_n = v_n{}^o\,(1 + l_1c_{1,n} + l_2c_{2,n} + \ldots.) \tag{8}$$

Then it is possible to calculate n, the number of carbon atoms in the component, by EQUATION 7 and, thus, a qualitative analysis is achieved.

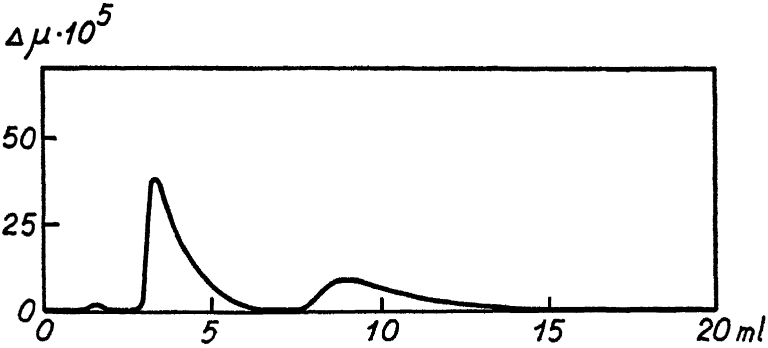


FIGURE 4. Elution analysis of 0.5 ml. of a solution containing 5 mg. each of lauric and palmitic acids. Filter: 500π carboraffin CI. Solvent (also used for elution) absolute ethanol. The small peak to the left is due to water in the ethanol.

Elution analysis. This procedure is identical with that of ordinary chromatographic analysis. After having washed the filter with pure solvent, a small volume of the solution to be analyzed is forced into the top of the column. The vessel A is then filled with pure solvent which is allowed to pass through the filter. Thus, the concentration of the solute in the cuvette C is zero during the course of the experiment, except when the different zones are moving through it. The resulting diagrams are shaped as in FIGURE 4, where each component is represented by a peak. The area A under a peak gives the amount of substance according to the equation

---

* It can, in fact, be shown that the simple EQUATION 6 is obtained if the equation of the adsorption isotherm can be put in the form $a_i{}^o = k_ic_i.r(c_1,\ldots c_n)$, where r is an arbitrary function but, of course, the same for all the components. This condition is necessary and sufficient (Claesson[20]).

$$A = \int_{v_1}^{v_2} c\, dv. \tag{9}$$

A quantitative analysis of the mixture is thus obtained by the simple measurements of areas. As every component comes out in a pure state, it can be identified by ordinary analytical methods. This means that the elution process is well adapted for preparative purposes, an advantage over the frontal analysis where only the weakest adsorbable component can be collected in a pure state. However, we have the disadvantage of the long "tails" of the peaks representing strongly adsorbed substances, which renders their separation difficult and makes the determination of the area A uncertain.

Displacement Development. This very ingenious method for adsorption analysis was introduced by Tiselius.[8] It eliminates all the disadvantages of elution analysis and is performed as follows. In the same way as in elution analysis, a small volume of the solution to be

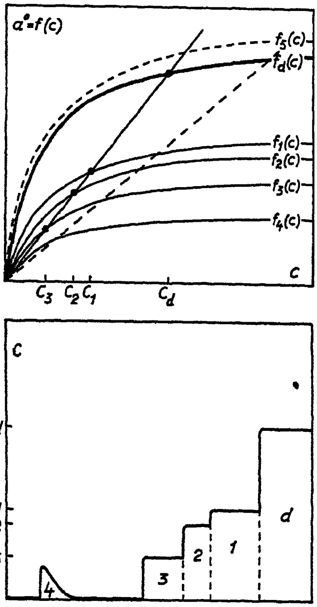


FIGURE 5. The construction of the displacement development diagram from the adsorption isotherms.

analyzed is forced into the top of the column, but instead of using pure solvent the column is afterwards washed with a solution of a substance, "developer", which is more strongly adsorbed than any of the components in the mixture. In fact, a frontal analysis of the developer is performed behind the mixture to be analyzed. By displacement, the components of the mixture are forced to move in front of the front of the developer. The components will also displace one another, so that they are arranged in order of increasing adsorption affinity, the one with the strongest adsorption affinity moving as a band immediately in front of the front of the developer (FIGURE 5). All the different components are thus obtained in pure state and the method is, therefore, also well suited for preparative purposes.

It is easily seen that, as soon as a stationary state has been established, all the zones in the column move forward with the same speed as the front of the developer. The latter has a speed which depends on its specific retention volume v_d^o. As each of the zones contains only a single component, we can apply EQUATION 2a and get

$$\frac{f_d(c_d)}{c_d} = \frac{f_1(c_1)}{c_1} = \frac{f_2(c_2)}{c_2} = \ldots . = v_d^o \qquad (11)$$

If the concentration of the developer c_d is kept constant in all experiments, then EQUATION 10 will give the values of the concentrations of all the components 1, 2, 3, . . . in the diagram. EQUATION 10 can easily be solved graphically. In the diagram showing the isotherms (FIGURE 5), a straight line is drawn from the origin to the point on the isotherm of the developer corresponding to the concentration c_d. The equation of the line will be

$$a^o = f(c) = v_d^o . c, \qquad (11)$$

and the intersections of that line and the corresponding isotherms will define the equilibrium concentrations $c_1, c_2, c_3,$ . . . From the figure, it is also evident that, if one component, 4, is adsorbed so weakly that the isotherm does not intersect the straight line, no equilibrium concentration is obtained. The component moves so rapidly through the filter that it cannot be overtaken by the developer, and it appears in the diagram as a free peak. By increasing the concentration c_d sufficiently, component 4 will also be displaced (dashed line in FIGURE 5).

From the foregoing, it is evident that, if the concentration c_d of the developer is kept constant in all experiments, the height of a step is independent of the amount of substance present in the step and only dependent of the nature of the substance. This *specific height, h,* is determined once and for all and is then used for the qualitative identification of the substance. As the area under a step is proportional to the amount of substance present and the height at the step is constant, the length of the step must be proportional to the amount of substance. When the proportionality factor b (length of step per gram of substance, *specific length*) is determined once and for all, the quantitative composition will be obtained by dividing the length of the steps by *b*.

For displacement development, we have thus obtained the important result that (1) the qualitative analysis is obtained by measurements of the heights of the steps, and (2) the quantitative analysis is obtained by measurements of the lengths of the steps. Displacement development is, therefore, the best method for adsorption analysis, but, unfortunately, it is not always applicable. Certain substances are adsorbed irreversibly and can neither be displaced nor eliminated, which leaves frontal analysis as the only possible way.

THE APPARATUS AND EXPERIMENTAL ARRANGEMENTS

The most important problem in the construction of an apparatus for adsorption analysis is the choice of a suitable method for the continuous determination of the concentration. In order that the method may apply as generally as possible, it is necessary to measure a property of the solution independent of the specific properties of the solvent and the solute. For that reason, the determination of refractive index was selected as most suitable. It is quite clear, however, that, in many special cases, a more specific and sensitive method may be used with greater advantage, as it is hardly possible to use refractive index determinations for solutions with a concentration smaller than 0.01%.

It is quite evident that the method for adsorption analysis discussed in the preceding section is also applicable in the case of gaseous systems, provided that the gaseous mixture to be analyzed is mixed with an inert gas, *e.g.*, nitrogen, which acts as a solvent. Therefore, an apparatus for adsorption analysis of gases and vapors will also be described.

Apparatus for Liquids. Two pieces of apparatus will be described; one, an interferometer which gives very high accuracy and is used for analytical purposes; the other, a self-recording apparatus with slightly lower sensitivity, which is used for preparative experiments and for analytical work on moderately dilute solutions.

The filter for the adsorbent and the container for the solution are of the same type for both pieces of apparatus. The adsorbent is packed in cylindrical tubes of brass, gold-plated inside and nickel-plated outside. The adsorbent rests on filter paper supported by perforated disks of metal. Filters of many different sizes have been used, the smallest having an inside diameter of 4 mm. and a height of 12.5 mm., the largest a diameter of 40 mm. and a height of 100 mm. For classifying the filters, their volume in mm.^3 has been used. In order to obtain simple numbers, the factor π has not been taken into account. The two filters mentioned above are, consequently, called 50π and $40{,}000\pi$, respectively. Four filters (50π, 250π, 500π, and 1250π) are shown in FIGURE 6. The filters are attached to the container for the solution by means of a heavy nut.

FIGURE 6. Filters for adsorbent.

When water is used as a solvent, the container for the solution is an ordinary tube of glass (FIGURE 7a) with fittings for the filter and for the rubber tubing from a pressure tank (ca. 3 kg./cm.^{2}). This pressure gives the solution a suitable speed through the filter (ca. 0.5 ml./min.). When

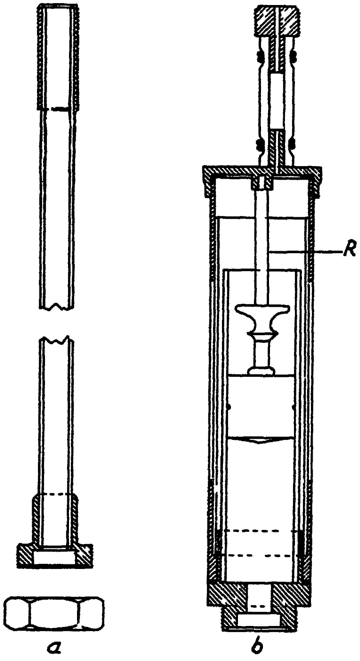


FIGURE 7. Containers for the solution.

organic solvents are used, it is necessary to prevent the air from the pressure tank being dissolved in the solution, as this would affect the refractive index. The air-free solution is therefore filled in a syringe (FIGURE 7b) enclosed in a metal tube. Then the air pressure acts on the piston and no air is dissolved in the solution.

The Interferometer. The interferometer is of the Rayleigh-Haber-Lowe type[21] and has been so designed that it is possible to use a very small volume of solution for the measurement. The cell consists of a cylindrical hole in a block of brass. The length of the hole is 80 mm and its diameter 1.4 mm. The filter with adsorbent is screwed into the cuvette and the solution flows through the cuvette and is then collected in a rack of graduated tubes where the volume of solution is measured. FIGURE 8 shows the cuvette mounted together with the syringe for the solution and a 1250π filter. There are four holes in the block: one for the solution, and the other three, reference cells filled with pure solvent. The cuvette is placed in a double thermostat. The light source is a tungsten ribbon lamp giving white light. The compensator is of the usual design, a plane plate of glass being turned by a micrometer

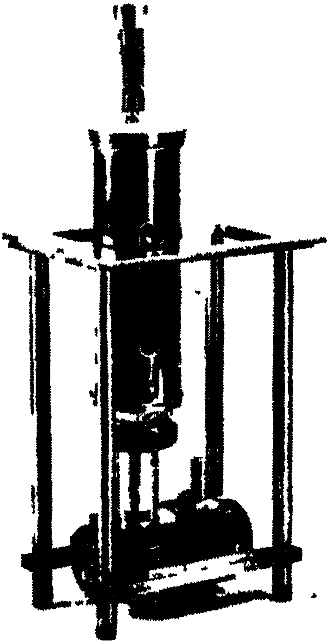


FIGURE 8. Interferometer cuvette, filter, syringe, and outlet tube mounted together.

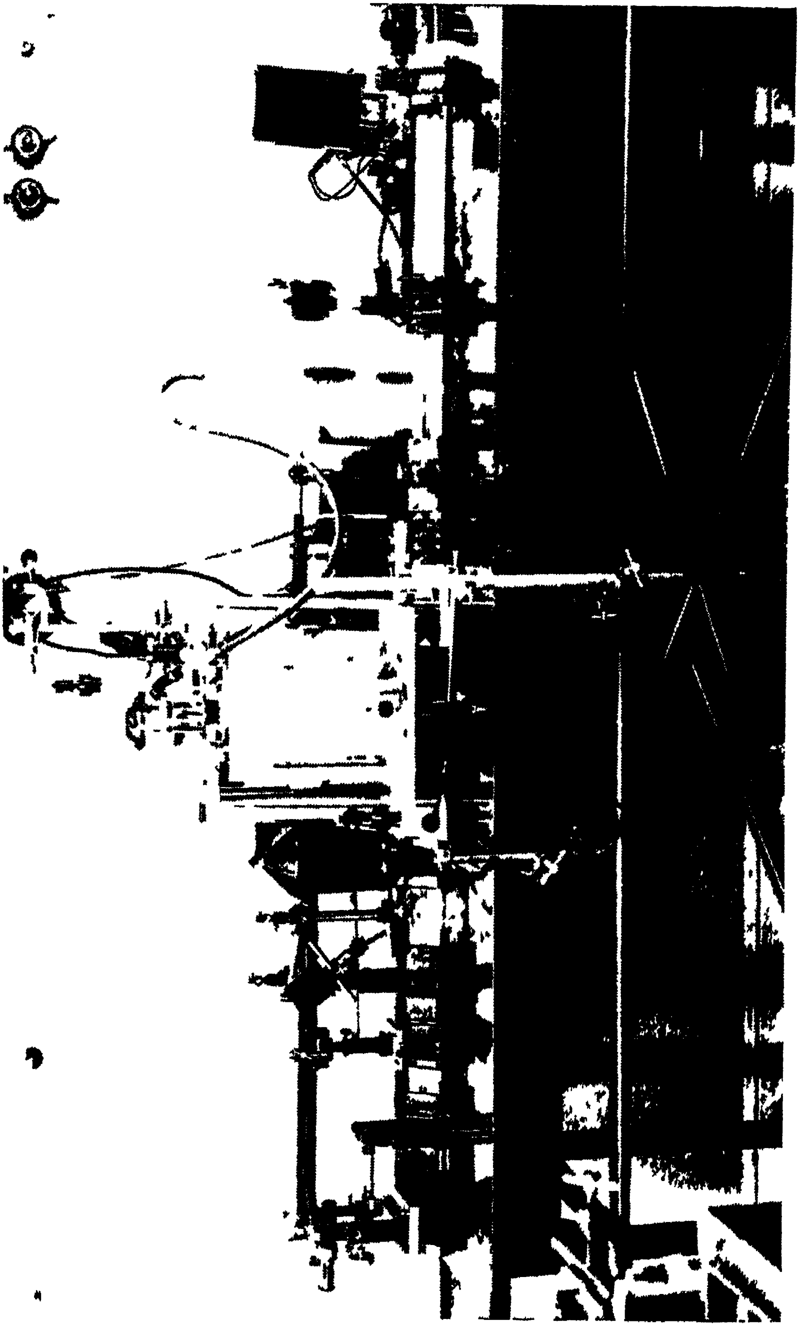


FIGURE 9. Photograph of the interferometric arrangement for adsorption analysis.

screw acting on a long arm. When measurements are performed on flowing solutions, the accuracy of reading amounts to 5.10^{-6} in refractive index and the total region of measurement is 6.10^{-3}. The values are obtained from the position of the micrometer screw, which has been calibrated once and for all by counting interference fringes in monochromatic light.

FIGURE 9 is a photograph of a complete apparatus. The rack with the test tubes is seen under a bell jar in front of the thermostat in the middle of the figure To the left are the ocular and two small telescopes, one being used for the reading of the micrometer screw, the other for the reading of the volume in the graduated test tubes.

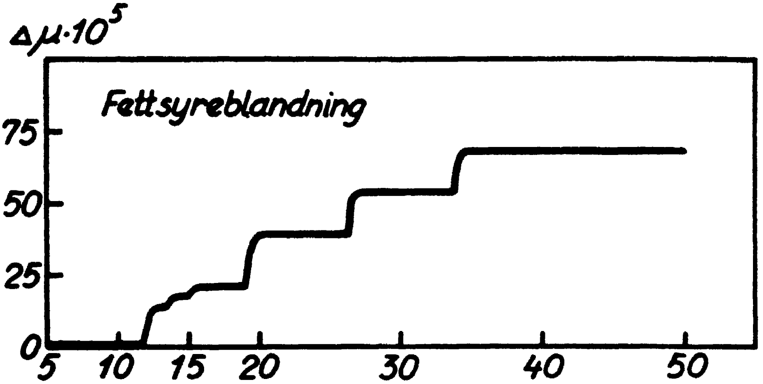


FIGURE 10 Frontal analysis of a mixture of six fatty acids (caprylic, pelargonic, capric, lauric, myristic, and palmitic acids). See TABLE 1.

FIGURE 10 shows a typical diagram obtained with this apparatus

The Self-Recording Apparatus. In this apparatus, the solution leaving the filter with adsorbent is passing one-half of a double hollow prism (D, FIGURE 11). The other half of the prism is filled with pure solvent. A beam of light passes the prism and makes a deflection which is proportional to the change in refractive index This deflection is very small (about 0.01 mm. if $l = 1000$ mm.) and is magnified in the following way. The beam of light is divided in two parts by a hexagonal prism of glass P, which fall on photoelements (C_1 and C_2). The two photoelements are connected by a galvanometer G in such a way that no current passes through it when the elements are equally illuminated. It can easily be shown that the galvanometer deflection is proportional to the change in refractive index in the prism The galvanometer is illuminated by the lamp L_3, and the light beam is reflected from the galvanometer mirror and the mirror M and makes a light spot on the photographic

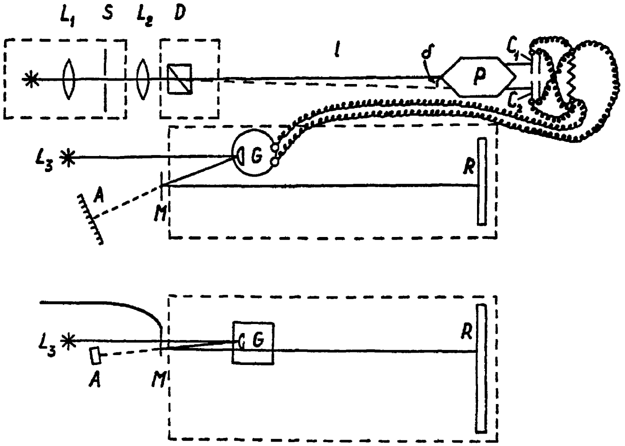


FIGURE 11 Sketch of the self-recording apparatus for adsorption analysis of solutions.

paper, *R*. When the refractive index changes, the light spot is consequently moved horizontally.

The volume of solution which has passed the cell is collected in a small flask hanging on a straight spring balance. The mirror *M* is attached to this spring and, consequently, the light spot moves vertically when the weight of the solution increases. It is thus seen that the refractive index of the solution and the weight of the solution which has passed the filter with adsorbent are recorded at right angles to each other on the photographic paper, *R* In this way, the desired diagram is obtained The whole apparatus is enclosed in a lightproof casing and is shown in FIGURE 12. The sensitivity is about 2.10^{-5} in refrac-

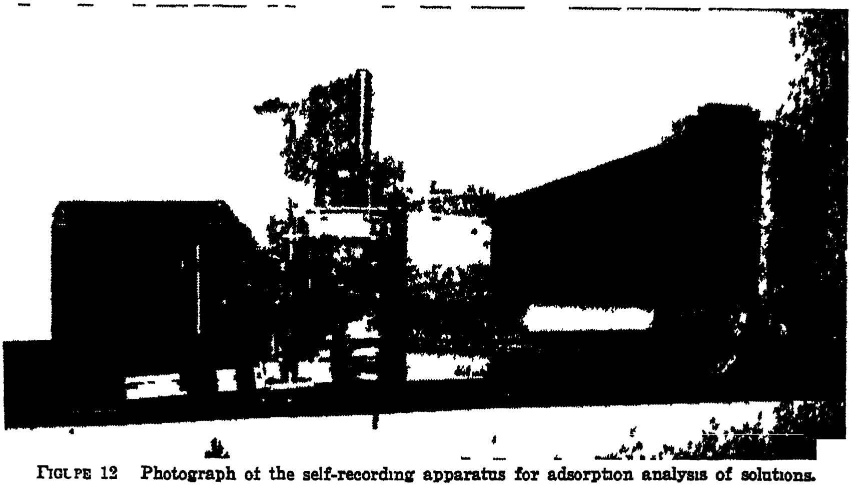


FIGURE 12 Photograph of the self-recording apparatus for adsorption analysis of solutions.

tive index, and, by using different springs, the apparatus can be used for volumes ranging from 10 to 1000 ml. The size of the photographic paper is about 20 by 30 cm.

PLATE 1a shows a typical curve obtained with this apparatus.

Apparatus for Gases and Vapors. The apparatus used for adsorption analysis of gases and vapors is shown in FIGURE 13. The mixture to be analyzed is mixed with a large amount of nitrogen ("the solvent") and introduced into the gas holder *A* through one of the glass tubes 25 or 26. The gas holder has been constructed according to Küster[22] and the compensating device 3-9 has the effect that the gas delivered always has the same pressure. From the gas holder, the gas passes the flowmeter *C*, the filter *D*, a cooling spiral immersed in water of 20°C, and the meter *F*, where the concentration is measured by observing the changes in thermal conductivity. *H* is a gas density balance, which can be used for special purposes. The meter *F* for the thermal conductivity contains four straight platinum wires mounted in four holes in a block of brass. The gas to be analyzed passes two of the wires, the other two being sur-

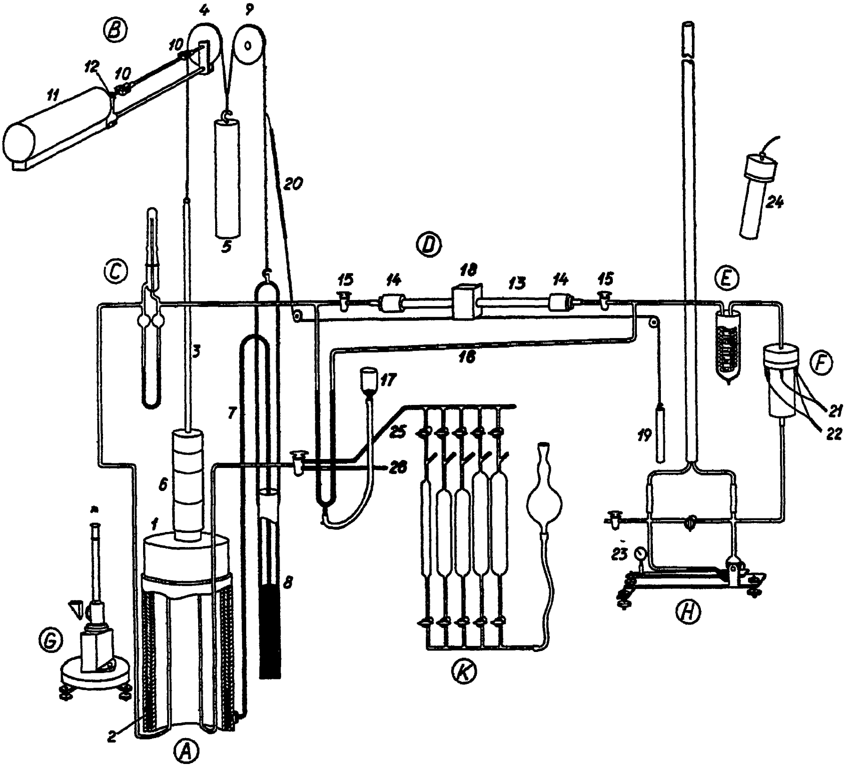


FIGURE 13. Sketch of the self-recording apparatus for adsorption analysis of gases and vapors.

rounded by a reference gas (*e.g.*, nitrogen). The four wires are heated electrically to about 125°C. and form the four resistances in a Wheatstone bridge. When the thermal conductivity of the flowing gas changes, the resistance of two of the wires changes, and the galvanometer *G* which is connected to the bridge, makes a deflection. This deflection is recorded on the drum 11, which is covered with photographic paper. It is seen from FIGURE 13 that the movement of the photographic paper is proportional to the movement of the gas holder and, consequently, proportional to the amount of gas that has passed the filter with adsorbent. The curve obtained thus has the desired shape with the concentration plotted against the volume.

PLATE 1b shows some typical curves obtained with this apparatus.

SOME EXPERIMENTAL RESULTS

The method for adsorption analysis described here has been applied to a large number of different types of substances by Tiselius and his co-workers.[1–20] Only a brief summary of these results can be given here, and the reader is referred to the original papers for more detailed information.

Some homologous series (fatty acids, ethyl esters of fatty acids, alcohols, dibasic acids) have been studied rather extensively by Claesson.[18] In this case, it was found that displacement development was not applicable and frontal analysis had to be used. With activated carbon as adsorbent and ethyl alcohol as solvent, good separation was obtained. It was also found that the theory developed for frontal analysis gave rather good results. The calculations were much simplified, as Traube's rule (EQUATION 7) was valid with good accuracy. This is seen from FIGURE 14 where $^{10}\log k$ has been plotted against the number, n, of carbon atoms in the fatty acids for some ethyl esters of fatty acids. A typical example of a frontal analysis of fatty acids has already been given in FIGURE 10. It is also seen there that the retention volumes for the first three acids (with 8, 9, and 10 carbon atoms) are more closely together than for the three following (with 12, 14, and 16 carbon atoms). This is in accordance with EQUATION 8. Some examples of quantitative analysis are given in TABLE 1, where the composition has been calculated by means of EQUATION 6.

It has also been found that branched and unsaturated fatty acids are adsorbed less than the corresponding normal saturated fatty acids when carbon is used as adsorbent. When silica is used as adsorbent and the solvent is a non-polar liquid, all the normal saturated fatty acids are adsorbed to almost the same degree, the branched acids are adsorbed less, and the unsaturated more. In that case, it is, consequently, possible to separate the fatty acid mixture into these three groups, and it has also

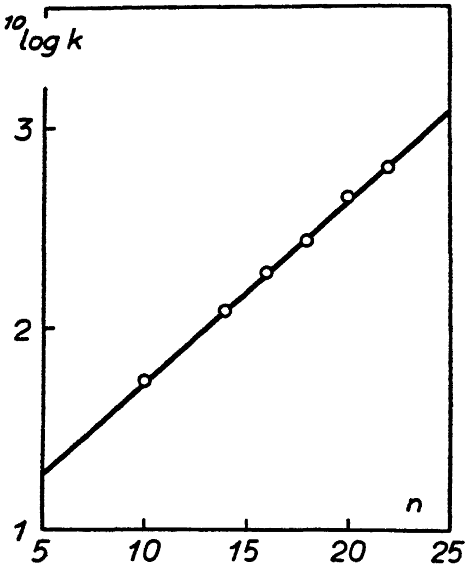


FIGURE 14. Traube's rule for ethyl esters of fatty acids.

been found that this can be done by means of displacement development (Claesson[19]). This effect is certainly due to the fact that the carboxyl group determines the adsorption on silica from non-polar solvents but not from polar solvents on activated carbon.

In some cases, it has been possible to use displacement development with extremely good results. One of the best examples of this is Tiselius' work on oligosaccharides (Tiselius and Hahn,[10] Weibull and Tiselius[15]). There, it was shown that a qualitative and quantitative analysis of mix-

PLATE 1

a. Frontal analysis of a mixture of 0.75% lauric and 0.75% palmitic acid in ethanol. Filter: 5000π "E-kol".

b. Displacement development of pure hydrocarbons. Developer: 1.5g. ethyl acetate. Adsorbent: 6g. carboraffin.

First row: 0.100 and 0.151g. *n*-pentane.
Second row: 0.151 and 0.200g. *n*-hexane.
Third row: 0.302 and 0.397g. *n*-heptane.

c. Displacement development of hydrocarbons. Developer: 1.5g. ethyl acetate. Adsorbent: 6g. carboraffin. Displaced substance:

0.137g. hexane + 0.143 heptane, 0.055g. hexane + 0.235 heptane, 0.018g. pentane + 0.104 hexane + 0.194g. heptane, 0.123g. pentane + 0.126g. hexane + 0.157g. heptane.

d. Frontal analysis of CCl_2F_2 containing a small amount of $CClF_3$.

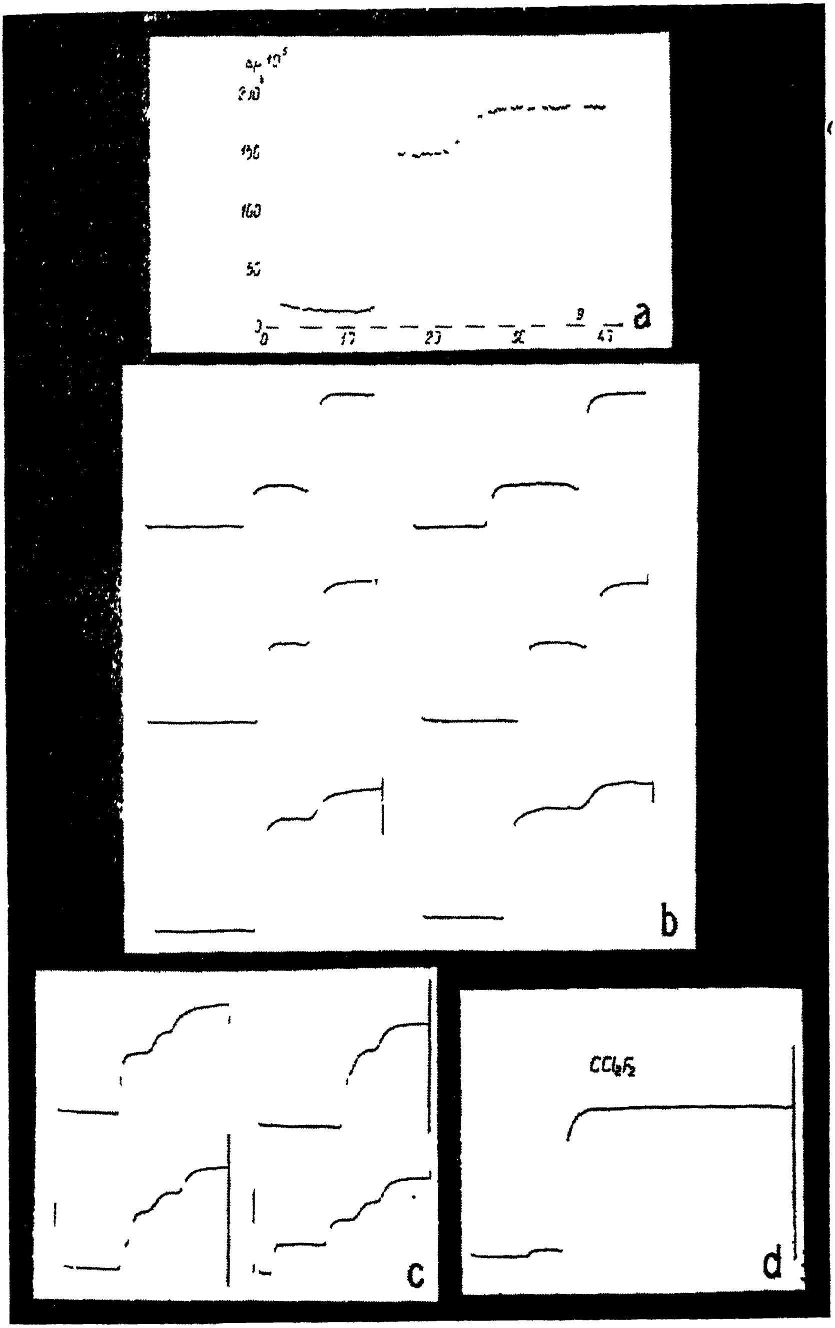


PLATE 1 (*see caption on opposite page*).

tures containing mono-, di-, tri-, tetra-, penta-, and hexasaccharides was obtained from the heights and lengths of the steps in the displacement development diagram. Values obtained by Tiselius for the retention volumes of some typical saccharides are given in TABLE 2.

TABLE 1

FRONTAL ANALYSES OF MIXTURES

Substances	Composition of the mixture (%)	
	Found value	Correct value
Caprylic acid, $C_7H_{15}COOH$	21	20
Capric acid, $C_9H_{19}COOH$	19	20
Myristic acid, $C_{13}H_{27}COOH$	20	20
Palmitic acid, $C_{15}H_{31}COOH$	41	40
Caprylic acid, $C_7H_{15}COOH$	14	10
Pelargonic acid, $C_8H_{17}COOH$	10	10
Capric acid, $C_9H_{19}COOH$	8	10
Lauric acid, $C_{11}H_{23}COOH$	22	20
Myristic acid, $C_{13}H_{27}COOH$	20	20
Palmitic acid, $C_{15}H_{31}COOH$	27	30
Adipic acid, $COOH(CH_2)_4COOH$	5	8
Suberic acid, $COOH(CH_2)_6COOH$	16	15
Sebacic acid, $COOH(CH_2)_8COOH$	31	31
Dodecanedioic acid, $COOH(CH_2)_{10}COOH$	48	46

TABLE 2

SPECIFIC RETENTION VOLUMES FOR SACCHARIDES
0.5 per cent solution in water, adsorbent Norit P3.

Substances	ml./g.
Arabinose	9.9
Xylose	14.8
Rhamnose	9.5
Levulose	13.6
Glucose	16.0
Galactose	17.3
Mannose	18.5
Saccharose	43.5
Lactose	51.0
Maltose	60.0
Raffinose	68.0

Tiselius[12, 17] has also described some preliminary experiments with amino acids and peptides.

When the self-recording apparatus is used for experiments with large amounts of substance, it is also necessary to use large filters. It is, then, often difficult to obtain sharp boundaries. The shape of the boundaries is very much improved if smaller filters are mounted below the larger filters, with a small chamber (ca. 2 mm. in height and of the same diameter as the larger filter) between the filters. In this way, a very marked sharpening of the fronts is obtained. Some experiments of this type are shown in FIGURE 15. The filters used there were 20000π, 5000π, and 1250π, mounted above each other. As the experiments were carried out by means of displacement development, the different components are obtained in a pure state. It is also seen, from the figure, that the heights of the steps are constant and the lengths of the steps are proportional to the amount of substance. The values for the amounts obtained from the diagrams agree within 3 per cent with the correct values. It is consequently possible to prepare substances in pure state, in that way, in quantities which amount to several grams.

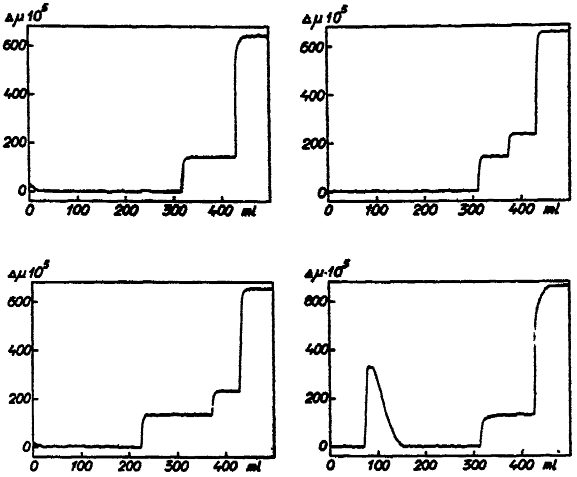


FIGURE 15. Displacement development of oligosaccharides. Developer: 4% phenol in water. Filter: 40000π + 5000π + 1250π carboraffin C. Displaced substance: 1.0 g. sucrose, 0.7 g. + 1.4 g. raffinose, 1.4 g. sucrose + 1.4 g. raffinose, 1.0 g. glucose + 1.0 = g. raffinose.

The apparatus for adsorption analysis of gases and vapors has until now mainly been used for analysis of hydrocarbons. It was found that displacement development worked well with these substances, and that esters of fatty acids were good developers. In PLATE 1b some experiments were shown, and in PLATE 1c it is seen that the same specific heights appear in the case of mixtures. From the great difference in specific height for pentane, hexane, and heptane, it is seen that the method is very selective and closely related compounds may be separated in this way. Some typical analyses of this kind are given in TABLE 3. It is seen that the accuracy of the quantitative analysis is good and the substances are well characterized by the heights of the steps. This height is given

TABLE 3

DISPLACEMENT DEVELOPMENT OF MIXTURES OF HYDROCARBONS
Developer: 1.5 g ethyl acetate; adsorbent: 6 g carboraffin

Substances	Height of step	Composition of the mixture (%)	
		Found value	Correct value
n-hexane *n*-heptane	56 78	20.4 79.6	19.7 80.3
n-hexane *n*-heptane	56 79	48.4 51.6	48.2 51.8
2-methylhexane *n*-heptane	65 77	40.0 60.0	40.7 59.3
n-pentane *n*-hexane *n*-heptane	32 57 80	5.6 33.0 61.4	5.7 32.8 61.5
n-pentane *n*-hexane *n*-heptane	30 57 81	30.2 31.1 38.7	30.1 31.9 38.0
n-hexane Benzene	56 89	43.1 56.9	43.2 56.8
Cyclohexane Benzene	67 97	46.4 53.6	46.6 53.4

as a percentage of the height of the developer, and this constant thus becomes independent of the special properties of the thermal conductivity meter.

Frontal analysis can, of course, also be used for analysis of gases. This is seen from PLATE 1d, which shows a frontal analysis diagram of $C\,Cl_2F_2$ containing a small amount of $C\,ClF_3$.

SUMMARY

In ordinary chromatographic analysis, colored bands are observed in a column filled with a suitable adsorbent. In the method described here, the concentration of the solution leaving the column is determined continuously in a small cell. This implies two fundamental advantages over the older chromatographic method. The procedure is especially suited for analysis of colorless solutions and independently of the color of the adsorbent, and, furthermore, the separation of the different components is greater in the percolate than in the column, which means increased selectivity. The theory for the method is first briefly discussed and, thereafter, the experimental arrangement is described. Two pieces of apparatus are described which are used for analysis of solutions, one being an interferometer, the other a self-recording apparatus. A self-

recording apparatus for adsorption analysis of gases and vapors is also described. Some typical examples of analyses carried out with this method are given in the last part of the paper.

BIBLIOGRAPHY

1. **Tiselius, A.**
   1940. Arkiv. Kem. Mineral Geol. 14B, **22.**
2. **Tiselius, A.**
   1941. Ibid. 14B, **32.**
3. **Tiselius, A.**
   1941. Ibid. 15B, **6.**
4. **Tiselius, A.**
   1941. Science **94:** 145.
5. **Tiselius, A.**
   1942. Advances in Colloid Science, ed. **E. O. Kraemer.** Interscience Publishers. New York.
6. **Claesson, S.**
   1941. Arkiv. Kem. Mineral Geol. 15A, **9.**
7. **Tiselius, A., & S. Claesson**
   1942. Ibid. 15B, **18.**
8. **Tiselius, A.**
   1943. Ibid. 16A, **18.**
9. **Tiselius, A.**
   1943. Kolloid Z. **105:** 101.
10. **Tiselius, A., & L. Hahn**
    1943. Ibid. **105:** 177.
11. **Claesson, S.**
    1944. 1884-1944. The Svedberg: 82. Almquist & Wicksell. Uppsala, Sweden.
12. **Tiselius, A.**
    1944. Ibid.: 370.
13. **Oehman, V.**
    1944. Ibid.: 413.
14. **Claesson, I., & S. Claesson**
    1944. Arkiv. Kem. Mineral Geol. 19A, **5.**
15. **Weibull, C., & A. Tiselius**
    1945. Ibid. 19A, **19.**
16. **Claesson, S.**
    1945. Ibid. 20A, **3.**
17. **Tiselius, A.**
    1947. Advances in Protein Chemistry. **III.**
18. **Claesson, S.**
    1946. Arkiv. Kem. Mineral Geol. 23A, **1.**
19. **Claesson, S.**
    1946. Rec. Trav. Chim. Pays-Bas **65:** 571.
20. **Claesson, S.**
    Arkiv. Kem. Mineral Geol. 24A.
21. **Haber, F., & F. Löwe**
    1910. Z. Angew. Chem. **23:** 1393.
22. **Küster, F. W.**
    1904. Z. Anorg. Chem. **42:** 453.

SOME EXPERIMENTS IN SYSTEMATIC QUANTITATIVE CHROMATOGRAPHY

By W. A. Schroeder

California Institute of Technology, Pasadena, California

During the past several years, our studies of the chemistry of propellants and explosives, under O.S.R.D. Contract OEMsr-881, have posed problems of the separation and quantitative determination of the substances in mixtures which contained from two to ten closely related or widely dissimilar organic compounds. Since any attempt to solve these problems by other than chromatographic methods would have been very difficult, we have consequently used chromatography for the quantitative determination of the compounds in these complex mixtures. I do not propose to describe the analysis of propellants and explosives as such, but rather to discuss the problems which arise in the adaptation of the usual chromatographic procedure to the analysis of mixtures of organic compounds with a precision comparable to that which is ordinarily obtained in inorganic analysis. These problems will be illustrated with data obtained at the California Institute of Technology by a group of about a dozen workers, to whom I wish to make full acknowledgment for the data which I shall present.

The materials which were analyzed in our studies of propellants and explosives may conveniently be classified in three categories: (1) mixtures of known qualitative composition; (2) mixtures whose composition could at least to some extent be surmised from their sources; and (3) mixtures of completely unknown composition. As an example of the first class, we may cite the quantitative determination of diphenylamine and diethyl phthalate in an American powder which contained, in addition, nitroglycerin and nitrocellulose. The second type of material is illustrated by various aged samples of this same powder, in which we wished to determine the fate of the stabilizer, diphenylamine. From information in the technical literature[1, 2] and from the chemical properties of diphenylamine, it was clear that provision must be made for the estimation of approximately twenty nitro and nitroso derivatives. Finally, explosives and powders of completely unknown composition were frequently encountered, and it was then often necessary to isolate and identify compounds which had not been studied previously, before a method for the complete quantitative analysis of the material could be devised.

Let us, first, turn our attention to the materials and methods which were used: the adsorbent, the developers, the elution, the determination of the compound in the eluate by spectrophotometry, as well as certain anomalous behaviors which were observed; and then let us consider the results of the application of chromatography to the quantitative analysis of the three types of materials which have been described.

All of the techniques which I shall discuss are based upon the use of silicic acid as an adsorbent. Silicious adsorbents have been used in a number of chromatographic studies.[3-11] Reagent silicic acid produced by several manufacturers has properties which recommend it as an adsorbent: it is pure white and thus permits ready detection of faintly colored zones; it is inert; it is finely powdered; and it may be modified in activity by appropriate treatment and yet, as received from the manufacturer, it does not show important variations from lot to lot. However, because of the slow filtration of solvents through the finely powdered adsorbent, it is advantageous to use a mixture of two parts by weight of silicic acid and one part of Celite 535.

For quantitative chromatography, it is important that adsorbent and solvents do not contribute impurities. Solvents may be purified by simple distillation in all-glass apparatus, and the adsorbent may be purified by pre-washing with redistilled solvents immediately before the chromatogram is run. Pre-washing of the adsorbent, however, has effects other than mere purification. These are illustrated in FIGURE 1. The variant in these experiments was the type of pre-wash, whereas the size of the column (19 x 150 mm.), the amount of 4-nitrodiphenylamine (0.5 mg.), and the type of developer (1:1 benzene-ligroin) remained constant. I shall speak of solvents in terms of the ratios of the volumes of the constituents and of the quantity of solvents in terms of the convenient convention "V ml." which we shall define as the volume of solvent which is required to wet completely a column of adsorbent. The alcohol pre-wash used V ml. of absolute alcohol and 2 V ml. of ligroin, the ether pre-wash

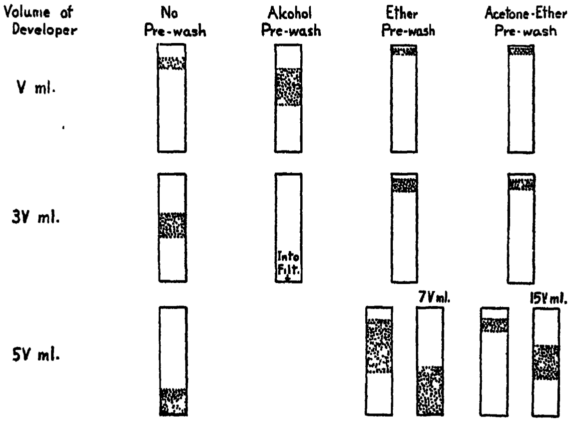


FIGURE 1. Effect of type of pre-wash upon the rate of development and width of a zone of 4-nitrodiphenylamine.

V ml. of ether and 2 V ml. of ligroin, and the acetone-ether pre-wash 0.2 V ml. of ether, V ml. of 1:1 acetone-ether, 0.8 V ml. of ether, and V ml. of ligroin. The zones as shown are idealized; the distribution of the compound was not uniform and the edges were not sharp.

A comparison of the position of the zone after equivalent development shows that the strength of the adsorbent increases in the following order: alcohol pre-washed adsorbent <unpre-washed adsorbent <ether pre-washed adsorbent <acetone-ether pre-washed adsorbent. Only the ether and acetone-ether pre-washes have been used in practice, and several differences may be noted in the behavior of the zone after these pre-washes. On ether pre-washed adsorbent, the zone moves with increasing speed and width in the lower sections of the column, whereas on acetone-ether pre-washed material, the rate of movement is regular and the increase in width, although apparent, is much less.

The explanation of the effect of the pre-wash is probably connected with the water content of the adsorbent, since, as is well known, the adsorptive strength of silica gel, of which silicic acid is surely a form, is a function of the water content and increases as the water content decreases until this content is decreased to about 5%.[12, 13] It has been found that, at 200° C., silicic acid loses weight for about 40 minutes, and that this adsorbent then behaves as though it had been pre-washed with acetone-ether. Another experiment showed that the loss in weight, on heating, is equivalent to the weight of the water which may be found in the solvents after pre-washing with acetone-ether. Apparently, then, the anhydrous solvents of the pre-wash activate the adsorbent by removing water. Ether, because it dissolves water less readily then acetone-ether, removes water less completely and less uniformly, with the consequent result that there is a gradation of the properties of the adsorbent from the top to the bottom of the column and, thus, the zones behave as indicated. Acetone-ether, however, produces a more uniform column, and the behavior of zones is more uniform as well. The de-activating effect of alcohol may be explained if we assume that alcohol, because of its structural similarity to water, merely replaces water in the adsorbent.

The adsorption affinity of many organic compounds on silicic acid is such that benzene or benzene-ligroin moves the zones at a reasonable rate but may not produce entirely satisfactory zones or separations. The pronounced effect of the developer upon the zone may be seen in FIGURE 2. The adsorbent was ether pre-washed, the only variant being the type of developer. The rates of development are similar, but the ether-ligroin produces a much more regular rate of movement and a more compact zone than does benzene-ligroin. Ethyl acetate or acetone in ligroin behaves very nearly as ether-ligroin.

Ligroin or benzene solutions of ether, ethyl acetate, or acetone, not only usually produce well-defined zones, but they often produce clean-cut separations of mixtures which benzene or benzene-ligroin can separate

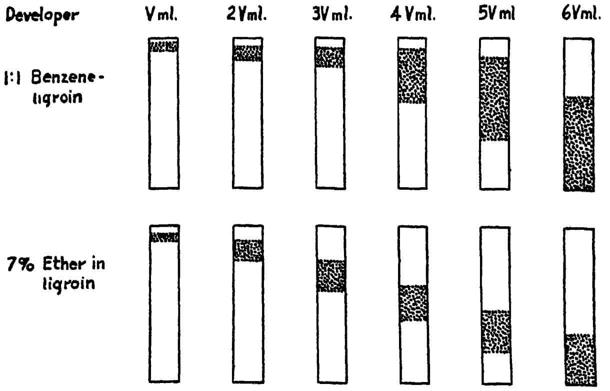


FIGURE 2. Effect of type of developer on the formation of a zone of 4-nitrodiphenylamine.

poorly if at all. However, developers which contain alcohol rarely are satisfactory. Sometimes, even ternary mixtures of benzene and ligroin with ether, ethyl acetate, or acetone are helpful. Three isomeric dinitrodiphenylamines may be used as an excellent illustration of the ease with which one developer will separate two compounds which cannot be separated by another. The relations may be seen in FIGURE 3. Thus,

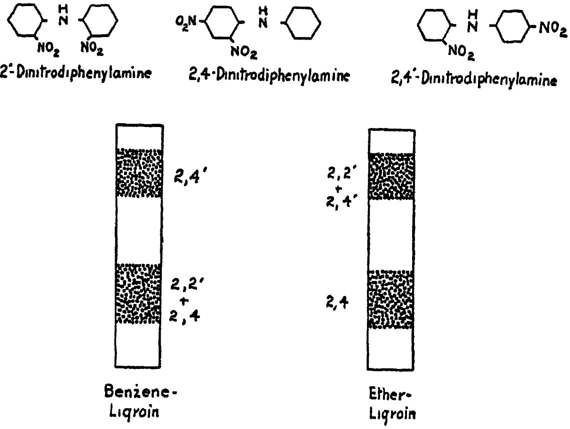


FIGURE 3. Separation of isomeric dinitrodiphenylamines.

benzene-ligroin readily separates 2,4'-dinitrodiphenylamine from the mixture of 2,2'- and 2,4-dinitrodiphenylamine, whereas ether-ligroin parts 2,4-dinitrodiphenylamine from the mixture of 2,2'- and 2,4'-dinitrodiphenylamine. The separation of the three compounds is, of course, relatively simple, since the mixture which may be isolated from one chromatogram may readily be separated by re-chromatographing with different development.

Common eluents such as alcohol, ether, and mixtures of alcohol or acetone in ether elute most compounds from silicic acid without difficulty. Indeed, only those compounds with an N-nitroso linkage required care in order to prevent decomposition. TABLE 1 shows typical results which may be obtained when separately weighed portions of a compound are chromatographed and eluted and the quantity is determined spectrophotometrically. Since the N,N-diphenylurea is so strongly adsorbed that ether itself is a good developer, it was necessary to use ethanol in ether as an eluent. It is seen that a 5 per cent solution was as effective an eluent as a 10 per cent solution. Usually, the calculated recovery will be 98 to 101 per cent of the weighed amount which was placed on the column. This precision is of the order which one might expect from a consideration of the errors involved.

TABLE 1

ELUTION OF N,N-DIPHENYLUREA FROM THE CHROMATOGRAPHIC COLUMN

Eluent % EtOH in Et_2O	Initial weight mg.	Weight recovered mg.	Per cent recovered
5	6.23	6.15	98.7
7.5	6.23	6.17	99.0
10	6.23	6.16	98.9
10	6.23	6.15	98.7
5	9.44	9.46	100.2
5	9.44	9.47	100.3
5	7.77	7.75	99.6

Since the spectrophotometer is so valuable in determining the quantities of 1 to 10 mg. which are most conveniently isolated, a few remarks about corrections in spectrophotometry should be made. Two familiar corrections are required in any optical instrument such as the Beckman Quartz Photoelectric Spectrophotometer which uses two optical cells, one for solvent and the other for solution. These corrections are necessary because few cells are exactly 1.000 cm. in thickness, and two cells are seldom so matched or so clean that one has no extinction relative to the other. Another correction which must be applied to spectrophotometric reading is one which corrects for small amounts of impurities which are not removed by pre-washing or distillation of solvents. This is estimated by means of a blank determination. The sum of the corrections rarely exceeds ± 1 or 2 per cent of the extinction. In an instrument of this type, the spectra of most compounds which contain benzene nuclei can be satisfactorily determined if the concentration is from 0.5 to 5 mg. per 100 ml., depending upon the intensity of absorption.

For definitive results by spectrophotometric methods, the solvent must be specified and, consequently, the eluent must often be replaced by a suitable solvent. Alcohol is an excellent solvent because its transparency permits readings at rather short wavelengths. However, it is in this transfer from one solvent to another that the greatest errors in a quantitative chromatographic determination may easily be introduced, since so much eluent must be removed from so little compound. The usual method of evaporating the solvent under reduced pressure is of course satisfactory if the eluted compound is non-volatile, but this is a very strict requirement, since the loss of 0.1 mg. would cause an appreciable error. On the other hand, solvent may often be removed from a volatile compound by allowing the solvent to evaporate from the zone after it has been cut from the column and then eluting the dry adsorbent directly with alcohol. Even a compound such as N-nitroso-N-ethylaniline, which is volatile and decomposes on prolonged contact with adsorbent, may be determined with reasonable accuracy if 100 mg. of a high-boiling compound such as diethylene glycol, which does not interfere with spectrophotometry, is added to the ether eluate before evaporation is begun. Thus, when the ether has evaporated, the 1 to 2 mg. of compound is dissolved in diethylene glycol and its vapor pressure correspondingly diminished. Actually, the recovery is only about 90 per cent, but without the high-boiling liquid it is about 10 per cent. It may even be that the solvents with which the column is wet and which, of course, are carried into the eluate by the eluent, do not interfere with spectrophotometry. For example, ligroin of boiling range 60-70° contains appre-

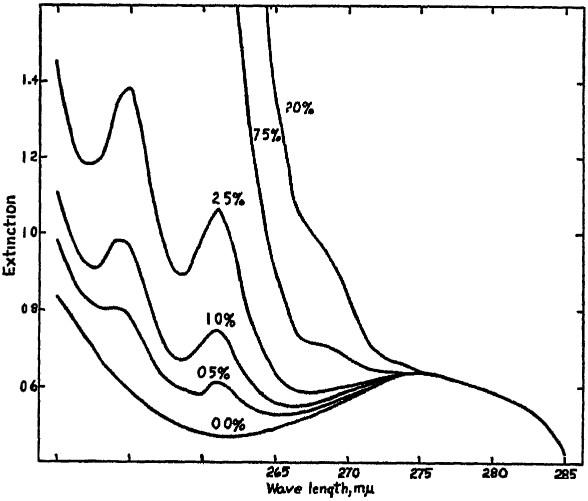


FIGURE 4. Extinctions of solutions containing the same concentration of diethyl phthalate and varying percentages of ligroin.

ciable amounts of benzene and, yet, 20 to 30 per cent of this solvent may be present in alcohol before the extinction differs appreciably from that of pure alcohol at wavelengths longer than 280 mμ. This effect is illustrated in FIGURE 4. The maxima at 255 and 262 mμ are caused by the presence of benzene. However, the benzene and ligroin have caused no increase in extinction at 275 mμ, the maximum of diethyl phthalate which is used for quantitative determinations. Thus, if the spectrophotometric constants of a compound are unaltered by the presence of the ligroin, measurement may be made directly on the alcohol eluate at wavelengths longer than 275 to 280 mμ, but at shorter wavelengths all extraneous solvents must be removed.

I should like to discuss, briefly, two anomalous behaviors which many compounds showed on silicic acid; they are inversions and what may be called "double zoning."

When one considers the number of chromatographic investigations which have been published in the last decade, comparatively few examples[6, 14-18] of inversion in the relative positions of two compounds by change of adsorbent or developer may be cited. On silicic acid, we have found that inversions by change of developer occur very frequently. A few examples which we have observed are shown in FIGURE 5. When a nitrate ester is one of a pair, an inversion is very likely to occur, because nitrate esters develop rather readily with benzene-ligroin mixtures and very much more slowly with ether-ligroin mixtures than would be expected, while most other compounds behave as one would anticipate. The examples in this figure, as well as others, have indicated that inverted compounds usually are rather dissimilar in structure. If we consider the dinitrodiphenylamines again (FIGURE 3), it is evident that, if the degree of change had been greater, an inversion would have occurred. Consequently, the ability or inability of one or another developer to separate

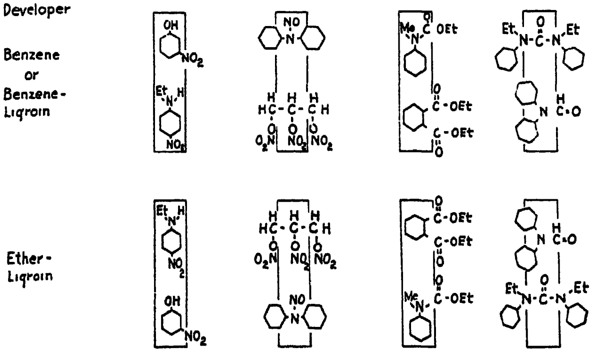


FIGURE 5. Inversion in relative position by changing developer.

two compounds, as well as inversions by change of developer, are merely degrees of the same effect. Some inversions, for example, that of N-nitrosodiphenylamine and nitroglycerin, have been used to advantage in devising separations. Thus, the separation of 2 mg. of N-nitrosodiphenylamine from 100 to 200 mg. of nitroglycerin with benzene-ligroin was not clean-cut because of the disproportionate amounts. Yet, when ether-ligroin was used, the nitroglycerin remained strongly adsorbed, while the N-nitrosodiphenylamine developed rapidly and separated easily. Although these instances of inversion are the result of rather extreme changes in the type of developer, examples might be cited of inversions which are caused merely by altering the proportions of the solvents in a developer. For example, N-ethylcarbanilide is more strongly adsorbed when a 2 per cent solution of ethyl acetate in ligroin is used, whereas 4-nitro-N-ethylaniline is more strongly adsorbed when the concentration is 5 per cent of ethyl acetate.

Of the approximately 100 compounds which we have chromatographed in studies of powder and explosives, a relatively large number exhibit the undesirable effect which may be termed "double zoning," that is, when a presumably homogeneous compound is chromatographed, it may, under certain conditions, produce two well-separated zones which cannot be distinguished by other than chromatographic methods, and each of which, when eluted and re-chromatographed, again produces two zones. Nearly always, the upper zone is the smaller of the two. Double zoning has been observed with such structurally different compounds as 4-nitroso-N,N-diethylaniline, N,N-diphenylurethane, carbanilide, and various nitro derivatives of diphenylamine. Although reasonable explanations of the cause may sometimes be given, no one explanation will suffice for all examples. In some instances, the presence of benzene in the developer seems to produce double zones which may, however, be prevented by changing the ratio of benzene to the other solvent in the developer; in other instances, precipitation of the compound by a succeeding developer in which it is insoluble may be the important factor. Double zoning appears to be similar to the blurred separation of the carotenoid methylbixin as observed by Zechmeister and Escue.[10]

Having described the methods and materials which were used in quantitative analyses, let us now consider the application of these methods to the analysis of the three types of material which have been mentioned.

The procedures which were required for the analysis of the powder of qualitatively known composition necessitated the separation of diphenylamine and diethyl phthalate not only from each other but also from nitroglycerin and nitrocellulose. Since diphenylamine, diethyl phthalate, and nitroglycerin may readily be separated from nitrocellulose by extracting the finely divided powder for several hours with methylene chloride in a Soxhlet apparatus, it is necessary to provide only for the chromatographic separation of diphenylamine, diethyl phthalate, and nitroglycerin in an extract of the powder. A study of the chromatographic properties of these compounds, individually and in mixture, showed that diethyl phthalate is considerably more strongly adsorbed than diphenylamine and nitroglycerin, which may easily be washed from the column, while diethyl phthalate remains strongly adsorbed. Diphenylamine, on the other hand, is the least strongly adsorbed and may be developed into the middle or lower portion of the column by developers which do not move diethyl phthalate and nitroglycerin from the top of the column. Thus, the actual chromatographic separation of the compounds presents no difficulties.

In addition to determining the method of separation, it was also necessary to consider the quantities which must be isolated. Since spectrophotometry was to be used, the spectrophotometric constants of the compounds were determined. The results showed that, if one milligram of diphenylamine was isolated, a dilution to 200 ml. with ethanol would give an appropriate value of the extinction, but that 15 to 20 mg. of diethyl phthalate would be required before the extinction would be sufficiently high at a satisfactory dilution. Since the approximate amount of diphenylamine in the powder was 0.7 per cent and that of diethyl phthalate 3.3 per cent, it may be calculated that 150-mg. and 500-mg. samples would be satisfactory for the respective analyses.

On the basis of this information, the following procedures were devised, as may be seen from FIGURE 6. About 2.5 g. of the powder were extracted

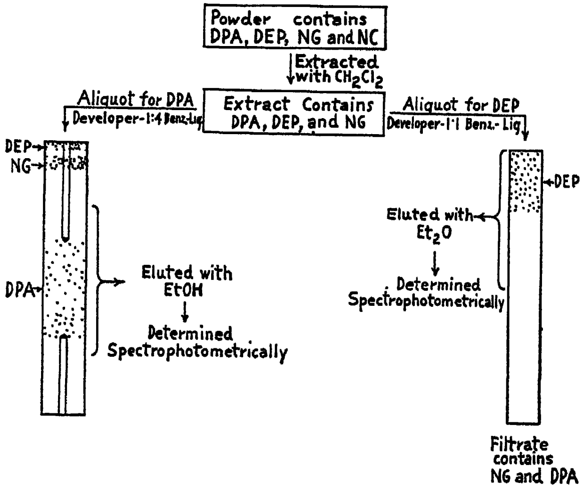

Powder contains DPA, DEP, NG and NC
Extracted with CH₂Cl₂
Aliquot for DPA
Developer-1:4 Benz.-Lig.
Extract Contains DPA, DEP, and NG
Aliquot for DEP
Developer-1:1 Benz.-Lig.
DEP
NG
DPA
Eluted with EtOH
Determined Spectrophotometrically
DEP
Eluted with Et₂O
Determined Spectrophotometrically
Filtrate contains NG and DPA

FIGURE 6. Determination of diphenylamine and diethyl phthalate.

with methylene chloride, which is equivalent to benzene as a developer. The extract in about 50 ml. of methylene chloride was then diluted to 100 ml. with ligroin, and 5-ml. aliquot portions were taken for the analysis for diphenylamine and 20-ml. aliquot portions for diethyl phthalate. The aliquot portions for diphenylamine were diluted with ligroin, so that the solvent was 1:4 methylene chloride-ligroin, whereas the 1:1 methylene chloride-ligroin was a suitable solvent for diethyl phthalate.

The diphenylamine was determined by chromatographing on a 25 x 150 mm. column of a mixture of 4:1 silicic acid-Celite which had been pre-washed with ether and ligroin. The diphenylamine was developed into the middle of the column with 1:4 benzene-ligroin and a post-wash, while the nitroglycerin and diethyl phthalate remained at the top. All of these compounds are colorless. Consequently, the zone of diphenylamine was located by streaking with a solution of sodium nitrite in sulfuric acid, which produces a blue color at the position of the zone, and the zone of nitroglycerin by streaking with diphenylamine in sulfuric acid, which also produces a blue color. Diethyl phthalate is more strongly adsorbed than nitroglycerin. In order to prevent the inadmissible loss which would result from the removal of a streak which had passed through a zone, the following method was adopted. The column was streaked carefully upward from the bottom until the production of color indicated the bottom of the zone and, likewise, from the top until the upper boundary had been found. After the zone had thus been delimited, the removal of the streak before elution resulted in inappreciable loss of compound. This particular streak reagent, sodium nitrite in sulfuric acid, in common with numerous other reagents, is affected by the solvents with which the column is wet. Thus, on clean adsorbent which is wet with benzene, an orange-brown color is produced, while on adsorbent which is wet with ligroin of boiling range 28-38° the streak is colorless. Consequently, after development of the chromatogram, it is well to post-wash with ligroin, in order to obtain the maximum contrast between the normal color of the reagent on free adsorbent and that color which is produced as the zone is reached. The zone was then eluted directly with alcohol and the quantity was determined spectrophotometrically. Diphenylamine is too volatile to permit evaporation of ether from an eluate, although the solvent may be permitted to evaporate from the zone before elution. However, studies showed that the solvents with which the column is wet and which are removed by elution do not interfere with spectrophotometric determination.

The diethyl phthalate was isolated by chromatographing on a 19 x 150 mm. column of ether pre-washed adsorbent and washing nitroglycerin and diphenylamine into the filtrate with 1:1 benzene-ligroin. The upper half of the column was eluted with ether, the solvent was evaporated under reduced pressure, a few milliliters of redistilled alcohol were added and evaporated to remove traces of residual solvent, and the residue was taken up in alcohol and estimated spectrophotometrically. Since no streak reagent was available for diethyl phthalate, the column had to be cut arbitrarily. However, experiments showed that a generous margin of safety was present if the upper half of the column was eluted.

A long study of these procedures showed that, usually, the results from duplicate aliquot portions of the same extract agreed to within one per cent of the amount of the compound, and that the values from duplicate extracts of the same powder had a maximum spread of about 3 per cent of the compound present. Thus, if 3.30 per cent of diethyl phthalate were actually present, values might range from 3.25 to 3.35 per cent, or 0.1 per cent of the powder. Likewise, if 0.70 per cent of diphenylamine were present, the values might range from 0.69 to 0.71 per cent, or 0.02 per cent of the powder. These results are representative of those which we obtained with quantitative chromatographic-spectrophotometric procedures for other compounds.

I should mention that no chromatographic-spectrophotometric procedure for nitroglycerin was devised because the compound has no maximum in its spectrum at wavelengths between 220 mμ and the end of the visible range, and, furthermore, chemical means, such as reduction with ferrous chloride and subsequent titration with titanous chloride of the ferric ion which is formed, are satisfactory for its determination.

The actual study of the changes and reactions of the stabilizer, diphenylamine, in the aging powder was preceded by investigation of the chromatographic and spectrophotometric properties and methods for the separation of the approximately twenty anticipated derivatives. Subsequent application of the procedures to the various heated samples then gave the results which are shown in FIGURE 7. The thirteen compounds of which the increase and decrease are represented in this figure, as well as four others (2,4-dinitrodiphenylamine, N-nitroso-2-nitrodiphenylamine, N-nitroso-2,4′-dinitrodiphenylamine, and N-nitroso-4,4′-dinitrodiphenylamine) which were detected in only insignificant amount, are a complex system which includes nitroso, mononitro, nitrosonitro, dinitro, nitrosodinitro, trinitro, tetranitro, pentanitro, and hexanitro derivatives of diphenylamine. The decrease of diphenylamine is not shown in FIGURE 7. The original 7 mg. per g. decreased at an essentially constant rate and was exhausted in about 5 days. The diversity of structure has permitted a few generalizations about the effect of the number and position of the substituent groups upon the relative adsorption affinity of the compounds. Thus, diphenylamine itself is very weakly adsorbed, but 4-nitrodiphenylamine has a markedly increased adsorption affinity. On the other hand, 2-nitrodiphenylamine is only slightly more strongly adsorbed than diphenylamine. This effect of a nitro group in the 2 position is such that dinitro derivatives, with the exception of 4,4′-dinitrodiphenylamine, are all less strongly adsorbed than 4-nitrodiphenylamine, and the trinitro and tetranitro derivatives are less strongly adsorbed than 4,4′-dinitrodiphenylamine. Chelation of the 2-nitro group with the amino group may very well be the cause of this effect. More highly nitrated derivatives are very strongly adsorbed. An N-nitroso group may either increase or decrease the adsorption affinity: for example, N-nitrosodiphenylamine is more strongly adsorbed then diphenylamine, but N-nitroso-4-nitrodiphenylamine is less strongly adsorbed than 4-nitrodiphenylamine, although still more strongly adsorbed than N-nitrosodiphenylamine.

Because of the relationships in the adsorption affinities, the methods which were devised for separating this complex mixture were not unduly complicated. The weakly adsorbed compounds could be isolated easily by individual chromatograms. The more strongly adsorbed compounds could then be separated into several groups, and the compounds in the groups could be parted by re-chromatographing with an appropriately different developer. The most complex mixture was that in the 35-day sample which contained no less than 10 derivatives. Usually, the analyses were less difficult because of the smaller number of compounds which had to be isolated. In general, the complete analysis of one sample required 3 to 4 days and 10 to 12 chromatograms.

These methods were not as refined as those for diphenylamine and diethyl phthalate, but it is probable that these results indicate, to within 5 or 10 per cent, the content of the derivatives in the powder.

The derivatives which are shown in FIGURE 7 account for only 60 per cent of the original diphenylamine. The fate of the remainder is still unsolved and, in any event, is a problem beyond the scope of this discussion.

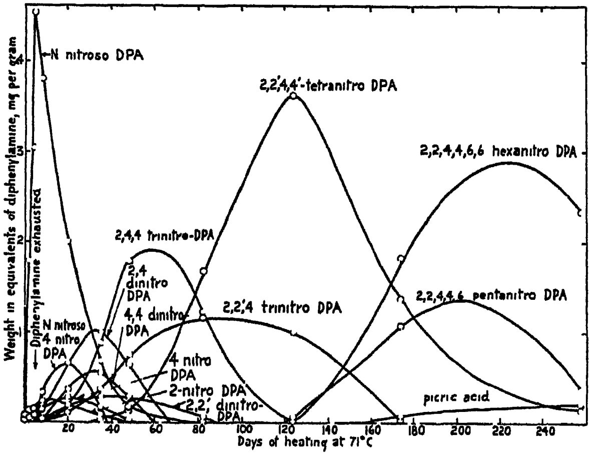


FIGURE 7. Derivatives formed from diphenylamine in double base powder.

Captured enemy munitions are perhaps the best examples of mixtures of unknown composition which it was our task to analyze. The general method of study, which is similar to the usual methods for, say, a natural extract, may be described, briefly, as follows. An extract of the powder was placed on a column and developed only slightly. Since the major constituents of most powders are colorless, it was necessary to detect the zones by application of various reagents which indicated whether a nitric ester such as nitroglycerin or diethylene glycol dinitrate as well as other compounds were present. The nitric ester was then tentatively identified by comparison of its chromatographic properties with those of known compounds. A few chromatograms with different developers usually indicated a method of separation for the other detected compounds, which were then shown by spectrophotometric examination to be known or unknown. Unknown compounds were isolated in sufficient quantity to permit identification by standard methods. Eluates of sections of the column which did not react with reagents were examined for the presence of undetected compounds. After the qualitative analysis was complete, a method for the quantitative analysis was devised and applied. In one instance, however, when the identification of one compound was difficult, its spectrophotometric properties were determined and the quantitative analysis was completed before the qualitative. TABLE 2 shows the results of the analysis of a German powder.

TABLE 2

COMPOSITION OF A GERMAN POWDER

Compound	Percentage
N-Ethyl-N-phenylurethane	1.58, 1.57
N,N-Diphenylurethane	1.75, 1.79
N,N-Diphenylurea	0.18, 0.19
Diethylene glycol dinitrate	32.7
Nitrocellulose and inorganic compounds	62.8
Total	99.0

The two results which are cited for the urethanes and diphenylurea were obtained from two separate extracts of the powder. Although the sum of the constituents is not 100 per cent, no evidence for the presence of other compounds was obtained, and there is every reason to believe that the discrepancy results from inaccuracies in the determination of diethylene glycol dinitrate and nitrocellulose, which are the major components and the most difficult to determine accurately. These were determined by standard methods, since no chromatographic-spectrophotometric method was devised for their determination.

The applications of quantitative chromatography which I have described have been in a specialized field, but I believe they indicate that, when relatively simple stable organic molecules are to be determined, complex mixtures of organic compounds can be analyzed with comparative ease.

BIBLIOGRAPHY

1. **Davis, T. L., & A. A. Ashdown**
1925. Ind. Eng. Chem. **17**: 674.
2. **Becker, F., & G. A. Hunold**
1938. Z. ges. Schiess- & Sprengstoffw. **33**: 213.
3. **Whitehorn, J. C.**
1934. J. Biol. Chem. **108**: 633.
4. **Reich, W. S.**
1939. Biochem. J. **33**: 1000.
5. **Weitz, E., & F. Schmidt**
1939. Ber. **72**: 1740.
6. **Cassidy, H. G.**
1940. J. Am. Chem. Soc. **62**: 3076.
7. **Kaufmann, H. P.**
1940. Angew. Chem. **53**: 98.
8. **Trappe, W.**
1940. Biochem. Z. **306**: 316.
9. **Mair, B. J., & A. F. Forziati**
1944. J. Res. Natl. Bur. Stand. **32**: 151, 165.
10. **Coleman, G. H., D. E. Rees, R. L. Sundberg, & C. M. McCloskey**
1945. J. Am. Chem. Soc. **67**: 381.
11. **Kirchner, J. G., A. N. Prater, & A. J. Haagen-Smit**
1946. Ind. Eng. Chem. Anal. Ed. **18**: 31.
12. **McGavack, J., & W. A. Patrick**
1920. J. Am. Chem. Soc. **42**: 946.
13. **Jones, D. C.**
1925. J. Phys. Chem. **29**: 326.
14. **Duschinsky, R., & E. Lederer**
1935. Bull. Soc. Chim. Biol. **17**: 1534.
15. **Erlenmeyer, H., & H. Dahn**
1939. Helv. Chim. Acta **22**: 1369.
16. **LeRosen, A. L.**
1942. J. Am. Chem. Soc. **64**: 1905.
17. **Strain, H. H., W. M. Manning, & G. Hardin**
1944. Biol. Bull. **86**: 169.
18. **Strain, H. H.**
1946. Ind. Eng. Chem. Anal. Ed. **18**: 605.
19. **Zechmeister, L., & R. B. Escue**
1944. J. Am. Chem. Soc. **66**: 322.

FRACTIONATION AND ANALYSIS OF HYDROCARBONS BY ADSORPTION

By BEVERIDGE J. MAIR
National Bureau of Standards, Washington, D. C.

This report[1] reviews two methods for separating hydrocarbons by adsorption which were developed in connection with the work of the American Petroleum Institute Research Project 6 at the National Bureau of Standards. These adsorption processes have been used primarily to separate mixtures of aromatic, paraffin, and naphthene (or cycloparaffin) hydrocarbons, as in the gasoline and kerosene fractions of petroleum, into two portions, one containing only aromatic hydrocarbons, the other only paraffin and naphthene hydrocarbons. This resolution of the mixture into two portions simplifies the problem of analysis, since an analytical distillation of the separated portions may be used to obtain much more information than is obtainable from an analytical distillation of the original mixture.

Description of the First Method. With the first method, the mixture of hydrocarbons is introduced into the top of a column containing an excess of solid adsorbent (silica gel), and when it has completely entered the adsorbent a low boiling paraffin hydrocarbon, such as *n*-pentane, is added in sufficient quantity to remove from the column the paraffin and naphthene hydrocarbons but not the aromatic hydrocarbons. The latter are then removed by adding an appropriate desorbing liquid such as methanol. The paraffins, naphthenes, and pentane are thus obtained as a mixture from which the pentane is easily removed by distillation. The aromatic hydrocarbons are obtained as a mixture with pentane and methanol. The methanol is easily removed by extraction with water, and the pentane is removed by distillation. Refractive indices are determined on the filtrate fractions and are used to distinguish the aromatic portion from the paraffin-naphthene portion. Silica gel adsorbent of 28 to 200 mesh size, with a natural rate of flow without application of pressure, is normally used in this method.

This method was used on a comparatively large scale in a Pyrex glass column, 10 cm. in diameter and 300 cm. in length, holding approximately 15 kg. of silica gel. This column was capable of separating from 500 to 1000 ml. of aromatic material.

This method of separation was tested on a known mixture of 17 pure hydrocarbons, the normal boiling points of which covered a range from 60° to 174° C. The mixture included all of the 5 normal paraffins from *n*-hexane through *n*-decane, the isoparaffin 2-methylpentane, the 4 normal alkylcyclohexanes from cyclohexane through *n*-propylcyclohexane, and all of the 7 possible aromatic hydrocarbons from benzene through isopropylbenzene. The separation of the aromatic hydrocarbons from the paraffins and naphthenes was quantitative within the limits of measurement, and their recovery was complete within the normal operating loss of material in processing (see reference 1, first item).

---

[1] Only a brief summary of this work is given here, since a detailed description is available in the following references:
Mair, B. J. & A. F. Forziati. J. Res. NBS 32: 165. 1944.
Mair, B. J. J. Res. NBS 34: 435. 1945.
Rossini, F. D., B. J. Mair, A. F. Forziati, A. G. Glasgow, Jr., & C. B. Willingham. Proc. Am. Petroleum Inst. 23 (8): 7. 1942; Oil Gas J. 41 (27): 106. 1942; Petroleum Refiner 21(11): 73. 1942.

Description of the Second Method. With the second method, the mixture of hydrocarbons is introduced into the top of a column of silica gel and, when the liquid level just reaches the top of the silica gel, a suitable desorbing liquid, such as ethyl or isopropyl alcohol, is added. The desorbing liquid forces the hydrocarbon portion down the column, during which passage the hydrocarbon portion is fractionated according to the adsorbability of the various components. The paraffin plus naphthene hydrocarbons issue from the column first, followed by the aromatic hydrocarbons and, finally, by alcohol.

On a small scale, this method is particularly suitable for the determination of aromatic content, the analysis being made by determining the fraction of the total volume constituted by the aromatic portion. Refractive indices are determined on the filtrate fractions, and are used to allocate, to the proper portions, the volumes of those fractions which occur at the "breaks" between the paraffin-naphthene and the aromatic portions and between the aromatic portion and alcohol.

The apparatus used for the analysis is made of Pyrex glass and consists essentially of a reservoir sealed to the column proper, which itself consists of two sections, an upper section, 2.2 cm. in diameter and 50 cm. in length, and a lower section, 1.0 cm. in diameter and 75 cm. in length. The column holds 160 g. of silica gel and will separate about 20 ml. of aromatic material. The silica gel is finer than that used with the first method and is of a nearly uniform particle size, about 60 per cent being between the 200 and 325 mesh sizes. Pressures of 5 to 8 lbs./sq. in., applied at the top of the reservoir from a source of inert gas under pressure, are used with material of the gasoline range, and somewhat higher pressures with material of the kerosene range. Usually, fractions of 5 or 10 ml. are collected, except in the neighborhood of the "breaks" between the portions, where 1-ml. fractions are taken.

Results of the analyses of synthetic mixtures with this method indicate that the aromatic content can be determined to within ±0.2 per cent.

The second method was used also for the separation, on a small scale, of monoolefin from paraffin and from aromatic hydrocarbons. At room temperature, considerable reaction of the olefins, as by polymerization, occurs. This polymerization is reduced but not entirely eliminated by sweeping the column with nitrogen prior to the introduction of the hydrocarbon, and by operating at lower temperatures.

Current work on this second method at the National Bureau of Standards includes the extension of the process to larger laboratory scale operations and the investigation of differences in the adsorbability of hydrocarbons of the same type due to differences in molecular size and structure.

STEREOCHEMISTRY AND CHROMATOGRAPHY

By L. Zechmeister

California Institute of Technology, Pasadena, California

It is one of the oldest theses in chemistry that all properties of a given substance are functions of its chemical structure. This fundamental idea is the starting-point for manifold efforts, all of which have attempted to locate within the framework of the molecule those sections which are made responsible for certain physical or chemical properties. A detailed study of the functional groups should not, however, distract the attention of the organic chemist from such characteristics of the compound which, possibly, are caused by the general shape of the molecule. The classic field for such investigations is stereochemistry, which does not consider structural changes.

A study of the dependence of adsorbabilities on the stereochemical configuration will probably acquire an increasing importance for the theoretical chemist. For the chromatographer, the problem of the dependence of the adsorption affinity on the structure finds its parallel in the problem of the dependence of the adsorption affinity on the spatial configuration.

The pertinent experimental data in this field may conveniently be divided into two parts. First, we shall consider those asymmetric compounds in which no double bond is responsible for the existence of stereoisomers; and, second, those in which *cis-trans*-isomerism is caused by the presence of double bonds.

Optical Antipodes and Diastereoisomers. The interesting question, "How far can racemates be resolved on the Twsett Column?" naturally arises in any consideration of compounds of the first type.

The idea that an optically active adsorbent may be used for this purpose can be traced back to an early paper of Willstätter (1904) who, two years before Tswett's first article on chromatography appeared, suggested that important information about the nature of the dyeing process might be obtained by investigating whether or not asymmetrical wool or silk fibers can be selectively dyed by one of the antipodes which are present in a racemic dye. Since no such dye was available at that time, Willstätter conducted model experiments with racemic alkaloids which can be easily adsorbed on protein fibers. The results were negative. His idea was later taken up by Porter and Ihrig (1923), who made use of some dyes which were prepared by coupling a diazo solution with racemic aminomandelic acid. Under favorable conditions, the antipodes were retained differently, and the remaining solution was optically active. However, these observations could not be confirmed by some other authors.

Recently, Martin and Kuhn (1941) passed an endless wool strip through a solution of racemic mandelic acid and observed (when a suitable temperature gradient was applied) that, as adsorption progressed, the solution acquired rotatory power.

Henderson and Rule (1936, 1939) must be credited with a chromatographic resolution of the racemate of *p*-phenylene-*bis*-iminocamphor, even if the operations were lengthy and required much adsorbent. For example, when 30 mg. of the compound was developed with petroleum ether on 6 kg. of lactose, the specific rotations + 485° and — 728° were observed instead of ± 1500°.

In inorganic chemistry, a partial resolution of racemic triethylenediamine chromichloride, (Cr en$_3$) Cl$_3$+3 ½ H$_2$O, on columns made of *d- or l*-quartz was achieved by Karagunis and Coumoulos (1938), who carried out a fractional elution with 85 per cent alcohol.

It follows, from some recent reports, that the well-known method of the resolution of a racemic acid by coupling it with an asymmetric alkaloid can be made more effective by a combination with chromatographic procedures. Hass, de Vries, and Jaffé (1943) observed that a partial resolution of brucine-*d,l*-mandelate takes place when this salt is developed with benzene on glucose. Furthermore, according to Fischgold and Ammon (1941), if *d,l*-mandelic acid and an optically active alkaloid are retained simultaneously on charcoal, the unadsorbed portion of the acid shows rotatory power.

That, in principle, a separation of diastereoisomers can be carried out on optically inactive adsorbents, may be illustrated by the following examples.

Jamison and Turner (1942) developed *l*-menthyl-*d,l*-mandelate with petroleum ether on alumina. The 15 cm. top section yielded, upon saponification, a mandelic acid with $[\alpha]_D = -18.2°$, but the corresponding figure for a fourth such section was +64°. Stoll and Hofmann (1938) succeeded in resolving an isomorphous mixture of *d*-isolysergic acid-*d*-isopropanolamide and *l*-isolysergic acid-*d*-isopropanolamide which could not be separated otherwise. When the mixture was developed with

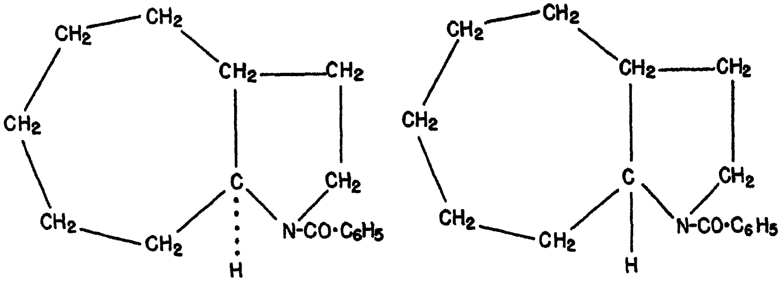


FIGURE 1. Diastereoisomeric N-benzol-cycloheptano-2,3-pyrrolidines.

acetone on alumina, the first few fractions from the liquid chromatogram yielded a crystalline *d*-derivative, $[\alpha]_D = +416°$; further portions showed first decreasing dextrorotation, then increasing levorotation, and the final fractions yielded practically pure *l*-compound, $[\alpha]_D = -342°$.

A clear separation of the two diastereoisomeric N-benzoyl-cycloheptano-2,3-pyrrolidines (FIGURE 1) was carried out by Prelog and Geyer (1945) by fractional washing on alumina with benzene and then with ether. First the α-isomer, m.p. 109°, then a mixture, and, finally, the pure β-isomer, m.p. 86-7°, passed into the filtrate.

Another recent contribution was made by Prelog and Wieland (1944). Previous to their work, it does not seem to have been possible to resolve a compound which contained an asymmetric trivalent nitrogen atom into its optically active components. These authors selected as starting material "Tröger's base" (FIGURE 2) and made use of petroleum ether on lactose in a liquid chromatogram procedure. The efficiency of this resolution is indicated by the fact that only 450 parts of sugar were necessary for 1 part of the compound, of which 5-6 per cent was resolved in a single run. After systematic recrystallizations of the highly rotating fractions, the values, $[\alpha]_D = +287°$ and $-278°$ were obtained.

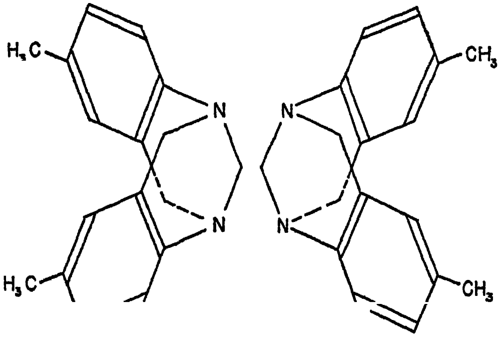


FIGURE 2. Tröger's base.

Prelog and Wieland's contribution is remarkable in two ways: first, by application of chromatography, the important stereochemical fact has been secured that, in structures like Tröger's base, an asymmetric trivalent nitrogen atom must occupy a stable position outside the plane of its three substituents; and, second, the possibility of a sequence inversion has been extended into the field of stereochemistry, since it was shown that, while a lactose column adsorbs the *l*-form of Tröger's base more strongly than it does the *d*-form, columns of *d*-tartaric acid or *d*-potassium hydrotartate prefer the *d*-isomer.

***Cis-trans* Isomerism around Nitrogen-Nitrogen or Carbon-Nitrogen Double Bonds.** It is well known that illumination of azobenzene solutions produces a lower melting isomer which is termed the *cis* form. It was shown by Cook (1938; Cook and Jones, 1939), and somewhat later by Zechmeister, Frehden, and Fischer-Jörgensen (1938) that the mixture of the two isomeric azobenzenes can be resolved on alumina by developing with benzene. The *cis* form is very well separated and occupies top position. The same sequence is exhibited by a number of azo dyes investigated by Cook. However, for *bis*-azo structures, like $C_6H_5.N{=}N.C_6H_4.N{=}N.C_6H_5$, he found the sequence from top to bottom to be *cis-trans, cis-cis,* and *trans-trans.*

The reversibility of the sequence of *cis*- and *trans*-azobenzene is implied by a study of Freundlich and Heller (1939), who showed that *cis*-azobenzene, which is the more hydrophilic isomer, is preferentially adsorbed by the more hydrophilic adsorbent, alumina. In contrast, the more hydrophobic *trans* form is the preferred one on the surface of a more hydrophobic adsorbent such as charcoal.

Examples of carbon-nitrogen double bonds which form *cis-trans* isomeric pairs are benzoin oxime, $C_6H_5.CHOH.C({=}NOH).C_6H_5$, and anisoin oxime, $C_6H_5.CHOH.C({=}NOH).C_6H_4.OCH_3$. Their *trans* forms, when developed on filtrol with benzene + 2% alcohol, are much more strongly adsorbed than the corresponding *cis* isomer. The stereoisomers can be located easily by brushing the column with ammoniacal copper sulfate. It should be mentioned that, in this instance, the *cis* configuration is not entirely stable under the conditions of the experiment and isomerizes on the column to give a few per cent of *trans* oxime (Zechmeister, McNeely, and Sólyom 1942).

***Cis-trans* Isomerism around Carbon-Carbon Double Bonds.** *Diphenylpolyenes.* In the field of the α,ω-diphenylpolyenes, $C_6H_5.(CH{=}CH)_n.C_6H_5$, there is some information available concerning the dependence of the adsorption affinities on the spatial configuration, especially for the members of this homologous series with $n = 1, 2,$ or 4. The usual synthetic methods produce *trans* forms; these, however, under the influence of illumination or other factors suffer a *trans* → *cis* rearrangement. Subsequent chromatographic separation shows that, at least for the three compounds mentioned, the all-*trans* form possesses a greater adsorption affinity than the corresponding *cis* isomer(s).

The location of the two stereoisomeric stilbenes ($n = 1$), and some of their derivatives, can be easily carried out after a development with petroleum ether, on alumina by means of permanganate (Zechmeister and McNeely, 1942).

Diphenylbutadiene ($n = 2$) was recently investigated (Sandoval and Zechmeister, 1947). In this case, also, upon illumination, two important characteristics are altered, *viz.,* the adsorbability decreases and the extinction values are reduced. At the same time, the initial fine structure shown by the spectral curve of the all-*trans* form disappears. Both the all-*trans* and a *cis* isomer can be located on the column by means of ultraviolet light, namely, by the intense fluorescence of the former and by the quenching of the weak fluorescence of alumina at the position of the latter.

In the case of diphenyloctatetraene ($n = 4$), a similar parallelism between adsorption and extinction changes exists (Zechmeister and LeRosen, 1942). The formation of most of the nine possible stereoisomeric forms of this compound is prevented by steric hindrance. The three fluorescing zones which can be seen on the calcium hydroxide column separate well from each other when developed with benzene-petroleum ether. The all-*trans* form occupies top position and is followed probably by *trans-cis-trans-trans-* and then by *trans-cis-cis-trans-*diphenyloctatetraene.

An interesting phenomenon was observed during chromatographic experiments with this polyene. When a *cis* zone was inspected for too long a time (say, for a whole minute) with the ultraviolet lamp while development was going on, it divided into two zones, one of which was all-*trans*-diphenyloctatetraene, newly formed on the column by photoisomerization. Thus, two all-*trans* zones were present on the same column, the second one at the place of its formation. This non-intended rearrangement took place on the cylindrical surface of the adsorbent, while the inside of the column showed only the original *cis* zone but none of a locally formed all-*trans* zone. In such cases, it is practical, before the solution is introduced, to cover the chromatographic tube with black paper except for a narrow observation slit.

Carotenoids. We are now going to discuss some of the services which have been rendered by chromatography in studies of the stereochemistry of the naturally occurring polyenes, the carotenoids. In this widely distributed class of pigments, the molecule is characterized by an unusual number of conjugated double bonds, for example, ten in α-carotene and eleven in β-carotene, γ-carotene, and lycopene ($C_{40}H_{56}$). For a time, it was believed by some authors that, within the framework of such a long conjugated system, each double bond loses so much of its individual character that no clearly defined *cis* and *trans* form can be expected.

As is well known, it was the chromatographic method which helped to disprove such concepts. The fundamental observation was made by Gillam and El Ridi (1936), who reported that carotene, when repeatedly chromatographed, produces a second zone, containing an isomer, below the unaltered main section on the Tswett column. The reversibility of the process was clearly recognized (Gillam and El Ridi, 1936; Gillam,

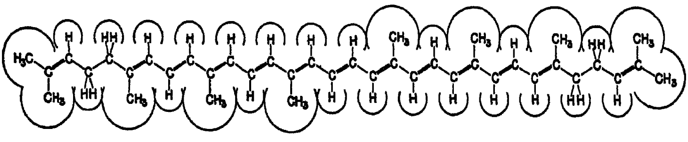


FIGURE 8. Model of all-*trans*-lycopene.

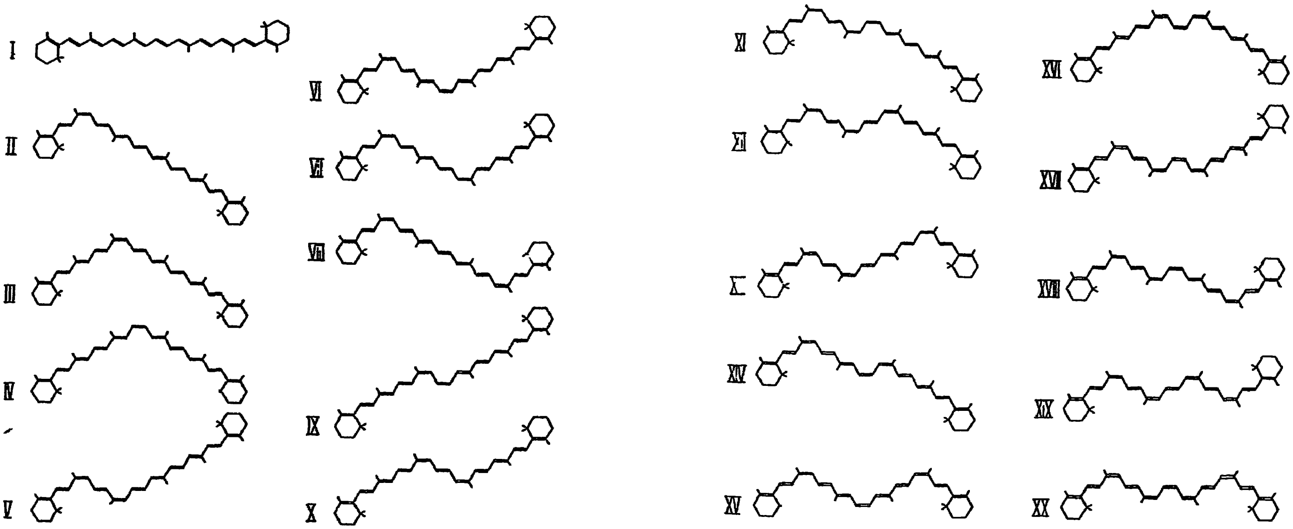


FIGURE 4. Skeleton models of the twenty members of the stereoisomeric β-carotene set.

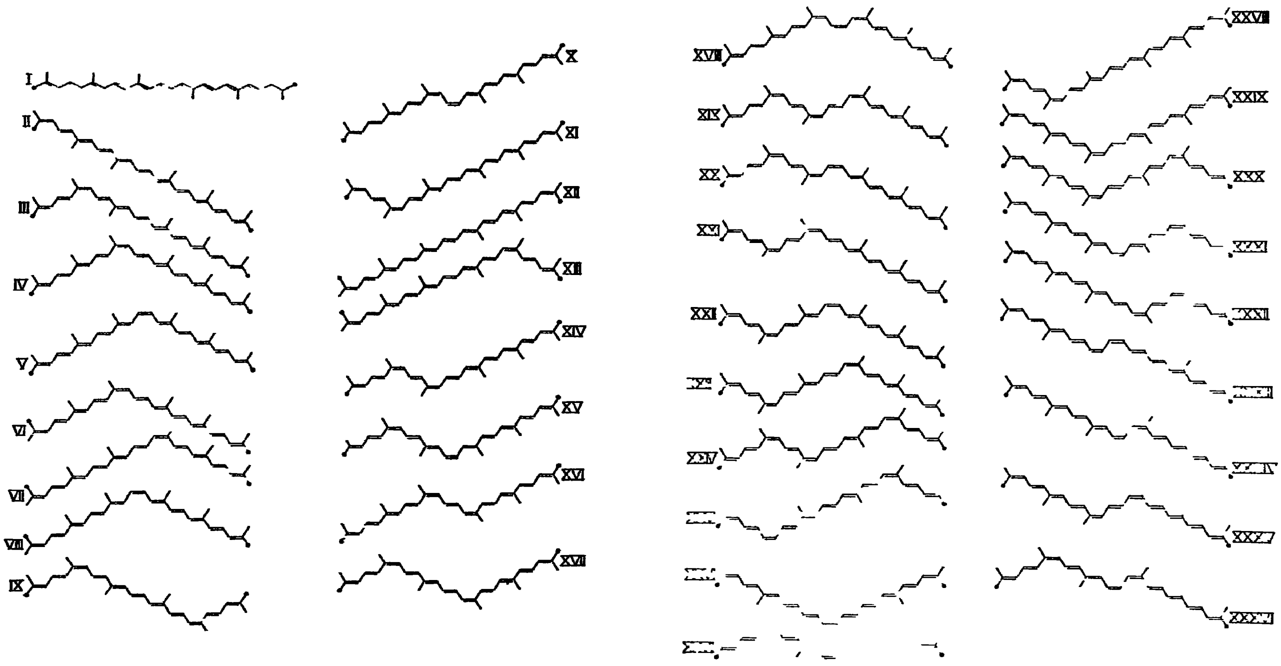


FIGURE 5. Skeleton models of the seventy-two sterically unhindered *cis-trans*-isomeric forms of lycopene (constructed by J. H. Pinckard): one all-*trans*-lycopene (I); four mono*cis*-lycopenes (II-V); twelve di*cis*-lycopenes (VI-XVII); nineteen tri*cis*-lycopenes (XVIII-XXXVI); nineteen tetra*cis*-lycopenes (XXXVII-LV); twelve penta*cis*-lycopenes (LVI-LXVII); four hexa*cis*-lycopenes (LXVIII-LXXI); and one hepta*cis*-lycopene (LXXII). The single side lines represent methyl groups; the black circles stand for the carbon skeleton of the end groups in the lycopene molecule.

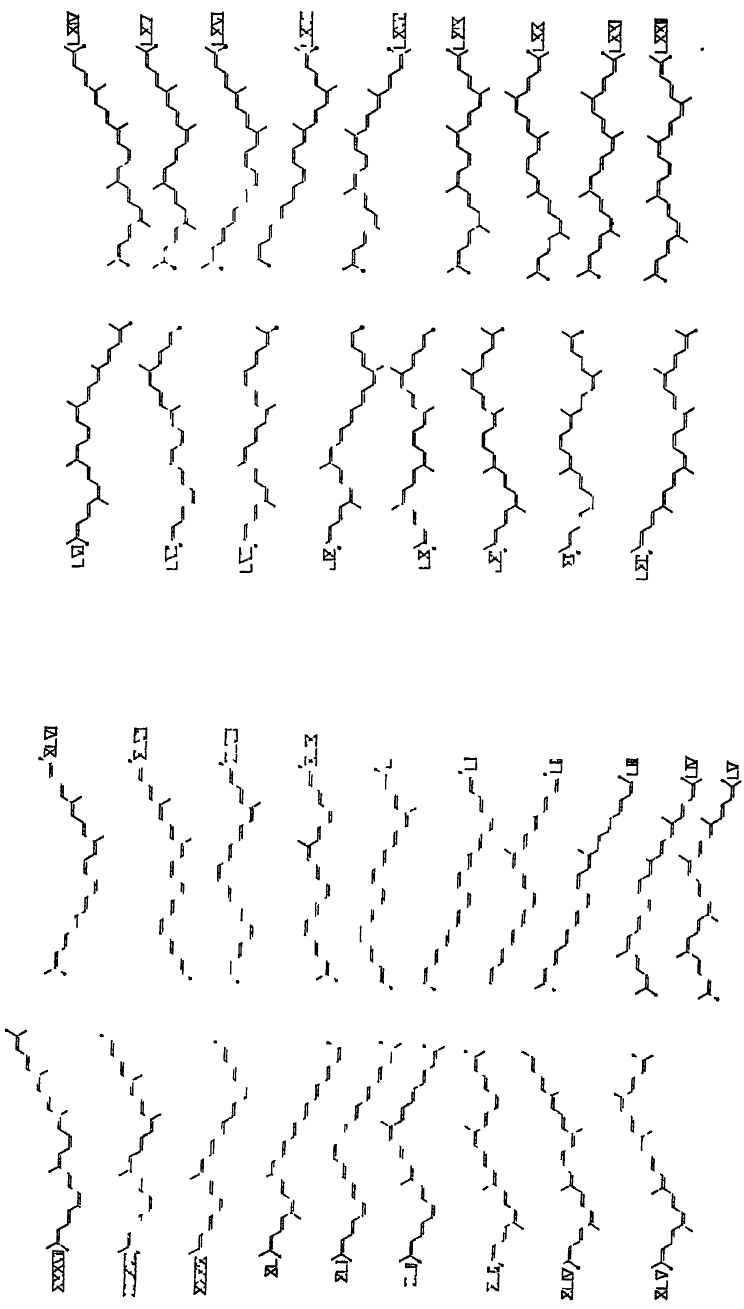


FIGURE 5 (*continued*).

El Ridi, and Kon, 1937). The isomerization of some carotenoids was also described, somewhat later, by the author and Cholnoky (1937), as well as by Strain (1938, 1941). It was shown by the author and Tuzson (1938) that this isomerization is independent of the adsorption process (confirmed by Carter and Gillam, 1939); and, furthermore, that the classic catalyst for *cis-trans* rearrangements, iodine, can be used also in the field of the C_{40}-carotenoids (1939; *cf.* Zscheile, Harper, and Nash, 1944. For a more detailed evaluation of the literature on isomerization methods, see a review written by the author, 1944; *cf.* Strain, 1938, 1939).

By far the major fraction of the naturally occurring carotenoid molecules is in the all-*trans* configuration (FIGURE 3), and even those equilibria or quasi-equilibria which can be produced by thermal, photochemical, or catalytic processes mostly contain the all-*trans* form as a main constituent and partially *cis* forms in lesser amount. The totality of all possible stereoisomers of a polyene is termed a "stereoisomeric set", and each isomer is a "member" of the set.

Calculation shows that, even for a relatively simple symmetrical compound such as β-carotene, the stereoisomeric set should include 272 members. However, according to Pauling (1939), four of the nine aliphatic double bonds cannot assume *cis* configuration, because of steric hindrance which is caused by the presence of methyl side chains. The expected number of stereoisomers is thus reduced to twenty (FIGURE 4). Even so, there is, evidently, no other method but chromatography which is applicable to the separation of stereoisomeric β-carotenes, of which about a dozen have been observed so far (Polgár and Zechmeister, 1942). The same statement is valid for the stereoisomeric forms of lycopene (FIGURE 5).

It is not appropriate, at this time, to discuss the methods which may be used to assign a tentative configuration to such isomers. A few remarks may therefore suffice (for a more detailed survey, *cf.* Zechmeister, 1944). Since each *trans* → *cis* rotation involves a shift in position of the spectral maxima to shorter wavelengths relative to the all-*trans* isomer,

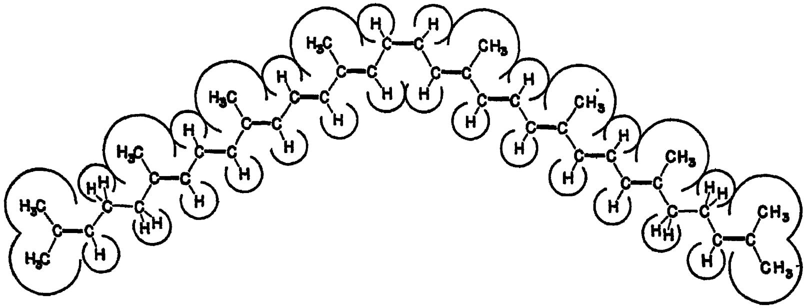


FIGURE 6. Model of central mono*cis*-lycopene.

the number of *cis* double bonds may be roughly determined. On the other hand, the so-called *cis*-peak effect, *i.e.*, the appearance of a new spectral maximum somewhere between 320 and 380 mμ (Zechmeister and Polgár, 1943) yields some information about the general shape of the molecule, that is, about the position of the *cis* double bonds. Those isomers possessing the highest *cis*-peaks are supposed to be V-shaped (FIGURE 6): in other words, a *cis* double bond is probably located at or near the center of the chromophore. In contrast, in isomers possessing low *cis*-peaks, the bending of the molecule has taken place at a peripheral section of the conjugated system (Zechmeister, LeRosen, Schroeder, Polgár, and Pauling, 1943).

Experiment has shown, in the case of carotenes, that bending of the molecule around a double bond may either increase or decrease the adsorption affinity. No parallelism is evident, in this case, between the direction of chromatographic and spectroscopic changes. That particular type of isomer, like neo-β-carotene U, which is rather stable and in which a centrally located *cis* double bond is assumed to be absent, possesses a higher adsorption affinity than the all-*trans* form, while the type of neo-B which shows a high *cis*-peak and greater liability is located below all-*trans*-β-carotene on the Tswett column.

The differences in adsorption affinity just mentioned are very considerable. In fact, they are of the same order of magnitude as effects which are caused by some structural alterations, *e.g.*, the addition or removal of a conjugated double bond. Thus, if a mixture of α- and β-carotene is catalyzed with iodine, in light, subsequent development with petroleum ether on calcium hydroxide demonstrates considerable overlapping of the two stereoisomeric sets, as shown by the following sequence (Zechmeister and Polgár, 1944):

(*top*) Neo-β-carotene U
Neo-β-carotene V
Neo-α-carotene U
All-*trans*-β-carotene
Neo-α-carotene V
Neo-β-carotene B
Neo-β-carotene E
Neo-α-carotene W
Neo-β-carotene F
All-*trans*-α-carotene
Neo-α-carotene B
(*bottom*) Neo-α-carotenes C, D, etc.

In connection with these observations, it was desirable to investigate how far the adsorption affinities are influenced by a great number of *cis* double bonds, and whether or not the rotation of, say, four to seven double bonds into *cis* position involves parallel or opposite changes in adsorbabilities and position of the spectral bands. Unfortunately, no method is available, at the present time, which permits the preparation *in vitro* of poly*cis* carotenoids. Evidently, the probability of the formation of such stereoisomers is very low under the experimental conditions applied. However, nature has provided us with the desired material in the form of several crystallized poly*cis* compounds, *e.g.*, prolycopene (FIGURE 7) and pro-γ-carotene, $C_{40}H_{56}$, which occur in several fruits (Zechmeister, LeRosen, Went, and Pauling, 1941; Zechmeister and Schroeder, 1942; LeRosen and Zechmeister, 1942; in petals: Schroeder, 1942). Recently, in collaboration with Pinckard (1947), six more poly*cis* lycopenes were observed which doubtlessly contain four to seven *cis* double bonds.

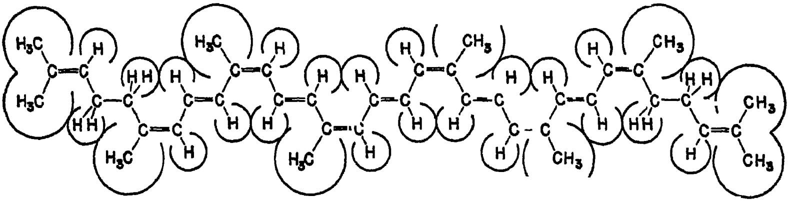


FIGURE 7. Model of prolycopene (a hexa*cis*-lycopene).

All of these compounds show a flat spectral curve in the *cis*-peak region, and a very much decreased adsorption affinity compared with that of the corresponding all-*trans* pigment or any stereoisomer containing a few *cis* double bonds as obtained *in vitro*. Evidently, when all or most of the sterically available double bonds are in *cis* configuration, the increasing or decreasing effect of a single *trans* → *cis* shift is overruled by the total effect of the poly*cis* configuration.

Our knowledge is less complete in the field of the dihydroxycarotenes, *e.g.*, lutein or zeaxanthin, $HO.C_{40}H_{54}.OH$, where no poly*cis* forms are available. As found by Strain (1938) and, independently, by our group (Zechmeister and Cholnoky, 1937; Zechmeister and Tuzson, 1939), all stereoisomers of the dihydroxy carotenes are adsorbed considerably above the all-*trans* form. This is also true for the important algal pigment fucoxanthin, $C_{40}H_{56}O_6$, as pointed out by Strain and Manning (1942).

Whether or not the increased adsorption affinity of *cis* xanthophylls is caused by the diminished distance between the two hydroxyl groups, cannot be decided at the present time. In favor of such an interpretation is the fact that the monohydroxy compound, cryptoxanthin, $HO.C_{40}H_{55}$, behaves like the hydrocarbons and yields stereoisomers with both increased and diminished adsorbability (Zechmeister and Lemmon, 1944).

The contrast in the behavior of a polyene hydrocarbon and the corre-

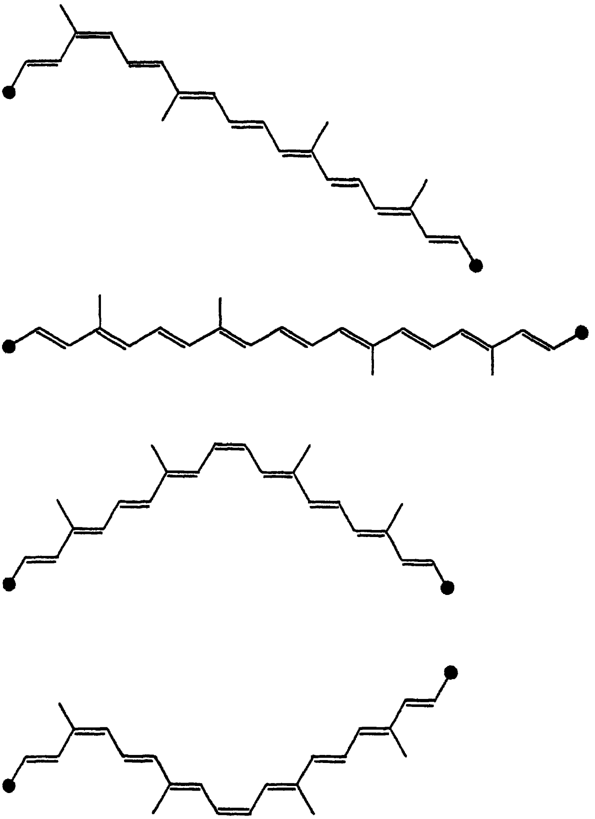


FIGURE 8. Skeleton models of four crystallizable stereoisomeric methylbixins, in the order of decreasing adsorption affinities (the dots designate —$COOCH_3$ groups): natural methylbixin (top); all-*trans*-methylbixin; neomethylbixin A; and neomethylbixin C (bottom).

sponding dihydroxy derivative can be illustrated, for example, by the following chromatographic sequences:

	α-Carotene Set:	*Dihydroxy-α-Carotene Set:*
(*top*)	peripherally bent isomers	V-shaped isomer
	all-*trans* form	peripherally bent isomer
(*bottom*)	V-shaped isomers	all-*trans* form

A lower molecular weight polyene, the dicarboxylic methylester, methylbixin, $CH_3OOC.C_{22}H_{26}.COOCH_3$, shows in principle the behavior of the carotenes, as illustrated by FIGURE 8 (Zechmeister and Escue, 1944).

We may summarize the available information in the field of *cis-trans* isomeric compounds containing sterically effective carbon-carbon double bonds as follows.

The adsorption affinities are highly dependent on the spatial configuration, often more so than on some structural features. If only one or two double bonds are present, classic procedures like crystallization may be used for the isolation of the stereoisomers. However, with a progressing lengthening of the unsaturated system and, thus, an increase of the number of the members of the set, chromatography becomes the only practical method for the resolution of a complicated stereoisomeric mixture. It seems even that the future development of certain chapters of stereochemistry will depend on the expert use of adsorption procedures.

A general theory of the dependence of the adsorption affinities on the spatial configuration is not available, and nothing is known concerning the anchoring points and the orientation of the individual stereoisomeric molecules on the active surface. On the other hand, the influence of *trans* → *cis* rearrangements on the chromatographic behavior is now better defined than previously, especially in the field of some natural polyene pigments.

Since the chromophoric group is responsible also for the spectrum, combined spectroscopic and chromatographic observations may be of considerable use in the future. Under the influence of spatial rearrangements, adsorbability and light extinction may undergo changes in a parallel or in an opposite direction. A closer study of such phenomena may offer an approach both for the interpretation of spectra and of adsorption processes. These problems are, however, of much deeper significance than those which could be discussed in the present short review of the relationship between stereochemistry and chromatography.

BIBLIOGRAPHY*

Carter, G. Ph., & A. E. Gillam
1939. The isomerization of carotenes. III. Reconsideration of the change β-carotene to pseudo-α-carotene. Biochem. J. **33**: 1325.

Cook, A. H.
1938. Preparation of some *cis* azo compounds. J. Chem. Soc.: 876.

Cook, A. H., & D. G. Jones
1939. *Cis*-azo compounds. J. Chem. Soc.: 1309.

Fischgold, H., & R. Ammon
1931. Contribution to asymmetric adsorption. Biochem. Z. **234**: 39.

Freundlich, H., & W. Heller
1939. The adsorption of *cis*- and *trans*-azobenzene. J. Am. Chem. Soc. **61**: 2228.

Gillam, A. S., & M. S. El Ridi
1936. The isomerization of carotenes by chromatographic adsorption. I. Pseudo-α-carotene. Biochem. J. **30**: 1735.

Gillam, A. E., M. S. El Ridi, & S. K. Kon
1937. The isomerization of carotenes by chromatographic adsorption. II. Neo-α-carotene. Biochem. J. **31**: 1605.

Hass, H. B., Th. de Vries, & H. H. Jaffé
1943. Methods for resolution enantiomorphs. III. Chromatographic adsorption. J. Am. Chem. Soc. **65**: 1486.

Henderson, G. M., & H. G. Rule
1936. A new method of resolving a racemic compound. Nature **141**: 917.
1939. New method of resolving a racemic compound. J. Chem. Soc.: 1568.

---

* This is not claimed to be exhaustive.

Jamison, M. M., & E. E. Turner
1942. The separation of diastereoisomerides by selective adsorption on optically inactive material. J. Chem. Soc.: 611.

Karagunis, G., & C. Coumoulos
1938. A new method for the separation of optical antipodes; its importance for optical activity. Nature **142**: 162.

LeRosen, A. L., & L. Zechmeister
1942. Prolycopene. J. Am. Chem. Soc. **64**: 1075.

Martin, H., & W. Kuhn
1941. A multiplication process for resolving racemates. Z. Elektrochem. **47**: 216.

Pauling, L.
1939. Recent work on the configuration and electronic structure of molecules; with some applications to natural products. Fortschritte d. Chem. org. Naturstoffe **3**: 203.

Polgár, A., & L. Zechmeister
1942. Isomerization of β-carotene. Isolation of a stereoisomer with increased adsorption affinity. J. Am. Chem. Soc. **64**: 1856.

Porter, C. W., & H. K. Ihrig
1923. Asymmetric dyes. J. Am. Chem. Soc. **45**: 1990.

Prelog, V., & U. Geyer
1945. On the diastereomeric cycloheptano-2,3-pyrrolidines. Helv. Chim. Acta **28**: 576.

Prelog, V., & P. Wieland
1944. On the splitting of Tröger's base into optical antipodes, a contribution to the stereochemistry of trivalent nitrogen. Helv. Chim. Acta **27**: 1127.

Sandoval, A., & L. Zechmeister
1947. Some spectroscopic changes connected with the stereoisomerization of diphenylbutadiene. J. Am. Chem. Soc. **69**: 553.

Schroeder, W. A.
1942. Formation of pro-carotenoids in "Monkey flowers" under some conditions. J. Am. Chem. Soc. **64**: 2510.

Stoll, A., & A. Hofmann
1938. Partial synthesis of ergobasin, a natural ergot alkaloid as well as of its optical antipode. Z. physiol. Chem. **251**: 155.

Strain, H. H.
1937. Eschscholtzxanthin: a new xanthophyll from the petals of the California poppy, *Eschscholtzia californica.* J. Biol. Chem. **123**: 425.
1938. Carotene. XI. Isolation and detection of α-carotene, and the carotenes of carrot roots and of butter. J. Biol. Chem. **127**: 191.
1938. Leaf Xanthophylls. Carnegie Institution. Washington.
1941. Isomerizations of polyene acids and carotenoids. Preparation of β-eleostearic and β-licanic acids. J. Am. Chem. Soc. **63**: 3448.

Strain, H. H., & W. M. Manning
1942. The occurrence and interconversion of various fucoxanthins. J. Am. Chem. Soc. **64**: 1235.

Willstätter, R.
1904. On an experiment concerning the theory of dyeing. Ber. Deutsch. Chem. Ges. **37**: 3758.

Zechmeister, L.
1944. *Cis-trans* isomerization and stereochemistry of carotenoids and diphenylpolyenes. Chem. Rev. **34**: 267.

Zechmeister, L., & L. Cholnoky
1937. Investigations on the Paprika pigment. X. Liebig's Ann. **530**: 291.

Zechmeister, L., & R. B. Escue
1944. A stereochemical study of methylbixin. J. Am. Chem. Soc. **66**: 322.

Zechmeister, L., O. Frehden, & P. Fischer-Jörgensen
1938. Chromatographic separation of *cis*- and *trans*-azobenzene. Naturwiss. **26:** 495.

Zechmeister, L., & R. M. Lemmon
1944. Contribution to the stereochemistry of cryptoxanthin and zeaxanthin. J. Am. Chem. Soc. **66:** 317.

Zechmeister, L., & A. L. LeRosen
1942. Stereoisomeric diphenyloctatetraenes. J. Am. Chem. Soc. **64:** 2755.

Zechmeister, L., A. L. LeRosen, W. A. Schroeder, A. Polgár, & L. Pauling
1943. Spectral characteristics and configuration of some stereoisomeric carotenoids including prolycopene and pro-γ-carotene. J. Am. Chem. Soc. **65:** 1940.

Zechmeister, L., A. L. LeRosen, F. W. Went, & L. Pauling
1941. Prolycopene, a naturally occurring stereoisomer of lycopene. Proc. Nat. Acad. Sci. **27:** 468.

Zechmeister, L., & W. N. McNeely
1942. Separation of *cis* and *trans* stilbenes by application of the chromatographic brush method. J. Am. Chem. Soc. **64:** 1919.

Zechmeister, L., W. H. McNeely, & G. Sólyom
1942. Chromatography of *cis* and *trans* benzoin and anisoin oximes with application of the brush method. J. Am. Chem. Soc. **64:** 1922.

Zechmeister, L., & J. H. Pinckard
1947. Some poly*cis* lycopenes occurring in *Pyracantha* berries. J. Am. Chem. Soc. **69:** 1930.

Zechmeister, L., & A. Polgár
1943. *Cis-trans* isomerization and spectral characteristics of carotenoids and some related compounds. J. Am. Chem. Soc. **65:** 1522.
1944. *Cis-trans* isomerization and *cis*-peak effect in the α-carotene set and in some other stereoisomeric sets. J. Am. Chem. Soc. **66:** 137.

Zechmeister, L., & W. A. Schroeder
1942. Pro-γ-carotene. J. Am. Chem. Soc. **64:** 1173.

Zechmeister, L., & P. Tuzson
1938. Spontaneous isomerization of lycopene. Nature **141:** 249.
1938. Isomerization of carotenoids. Biochem. J. **32:** 1305.
1939. Reversible isomerization of carotenoids by iodine catalysis. Ber. Deutsch. Chem. Ges. **72:** 1340.

Zscheile, F. P., R. H. Harper, & H. A. Nash
1944. Photochemical reaction of iodine with carotenoids. Arch. Biochem. **5:** 211.

CHROMATOGRAPHY IN THE STREPTOMYCIN PROBLEM

By Robert L. Peck

Research Laboratories, Merck and Co., Inc., Rahway, N. J.

Streptomycin has become increasingly important since its discovery was first reported, in 1944, by Schatz, Bugie, and Waksman.[1] The isolation of pure streptomycin and the application of streptomycin to clinical practice in less than three years after its discovery represent a significant achievement. To a considerable extent, this accomplishment is due to the development of chromatographic methods for purification of the active substance. The present paper is concerned with this phase of the streptomycin problem.

The object of the reported work on the chromatographic purification of streptomycin was essentially the preparation of pure material for degradative and other studies, rather than the careful investigation of chromatographic behavior. It will be understandable, therefore, that there remains much to be desired concerning knowledge of the chromatographic behavior of streptomycin.

The problem was the separation of a complex active substance which was soluble in water and in methanol, but nearly insoluble in the common organic solvents, from large amounts of inactive impurities possessing very similar solubility properties. Chromatography appeared to be the only practicable approach, since neither purification by extraction from water into organic solvents and back into water, nor direct crystallization of the active substance was possible. The solubility of streptomycin concentrates was such that the solvent for chromatography was necessarily strongly polar (water, methanol, or mixtures largely composed of these with miscible solvents). Examination of adsorption behavior led to the general choice of alumina as the adsorbent for columns. Darco G-60 carbon was also found applicable.

The colorless nature and similarity in properties of streptomycin and many of the impurities present in the concentrates made it difficult to follow the chromatographic purification of streptomycin on the columns by visual observation. Although colored salts of streptomycin have been made,[2, 3] their chromatography does not appear to have been investigated to any great extent. The "liquid chromatogram", discussed by Zechmeister and Cholnoky,[4] has been employed, therefore, for the chromatographic work thus far described in detail.

---

[1] Schatz, A., E. Bugie, & S. A. Waksman. Proc. Soc. Biol. & Med. 55: 66. 1944.

[2] Fried, J., & O. Wintersteiner. Science 101: 613. 1945.

[3] Kuehl, F. A., Jr., R. L. Peck, A. Walti, & K. Folkers. Science 102: 34. 1945.

[4] Zechmeister, L., & L. Cholnoky. The Principles and Practice of Chromatography: 76, 77. Second Edition. John Wiley & Sons, Inc. New York. 1941.

The liquid chromatogram, in principle, consists of a series of filtrates successively collected from a chromatographic column. In this type of procedure, part or all of the adsorbate is forced through the column into the filtrate. Examination of the arbitrary series of filtrate fractions taken gives the required information on fractionation.

Investigation of liquid chromatograms of streptomycin has been based on concentration of solids and concentration of activity in the filtrate fractions. Other variables, such as optical rotation and nitrogen content, could also be followed. Plotting these results against the filtrate volumes gives a graphic picture of the chromatogram.

A brief discussion of the source and chemical nature of streptomycin is given here, in order to furnish a basis for an understanding of the type of compound being purified by chromatography. Streptomycin is produced by the mold, *Streptomyces griseus*, grown under certain specified conditions.[1] The antibiotic is usually recovered[2, 8, 5, 6, 7] from filtered culture broths by adsorption on activated charcoal. The charcoal is eluted with acid alcohol and the antibiotic is obtained as a crude, amorphous, hygroscopic powder. This concentrate may be purified by chromatography to yield nearly pure streptomycin salts.

Streptomycin preparations from chromatographic purification have been converted to pure crystalline salts, including the reineckate,[2, 5] helianthate,[3] sulfate,[3] *p*-(2-hydroxy-1-naphthylazo)-benzenesulfonate,[3] and the calcium chloride double salt.[8] From investigations carried out on the pure salts, and on degradation products obtained from them, much of the chemical nature of streptomycin has become known. Streptomycin is a levorotatory base, $C_{21}H_{39}N_7O_{12}$, with a molecular weight of 582. The trihydrochloride shows $[\alpha]_D$ —86.7°. Streptomycin appears to be a glycoside composed of a 1,3-diguanido-2,4,5,6-tetrahydroxycyclohexane, termed streptidine,[9] and a disaccharide-like substance designated streptobiosamine.[10] Streptobiosamine, $C_{13}H_{23}NO_9$, is composed of N-methyl-l-glucosamine and a six-carbon sugar-like fragment, designated streptose[10] or streptonose.[11] The reducing group of the N-methyl-l-glucosamine fragment is bound in glycosidic linkage to streptose in streptobiosamine. There appears to be an aldehydic group in the streptose portion of intact streptomycin,[11] and also a C-methyl group.[12] Reaction of streptomycin

---

[5] Carter, H. E., R. K. Clark, Jr., S. R. Dickman, Y. H. Loo, P. S. Skell, & W. A. Strong. J. Biol. Chem. 160: 337. 1945.

[6] Le Page, G. A., & E. Campbell. J. Biol. Chem. 162: 163. 1946.

[7] Vander Brook, M. J., A. N. Wick, W. H. De Vries, R. Harris, & G. F. Cartland. J. Biol. Chem. 165: 463. 1946.

[8] Peck, R. L., N. G. Brink, F. A. Kuehl, Jr., E. H. Flynn, A. Walti, & K. Folkers. J. Am. Chem. Soc. 67: 1866. 1945.

[9] For leading references to the structure of streptidine, see: Peck, R. L., C. E. Hoffhine, Jr., R. P. Graber, F. W. Holly, R. Mozingo, & K. Folkers. J. Am. Chem. Soc. 68: 776. 1946.

[10] For leading references to the structure of streptobiosamine, see: Kuehl, F. A., Jr., E. H. Flynn, N. G. Brink, & K. Folkers. J. Am. Chem. Soc. 68: 2096. 1945; and Brink, N. G., F. A. Kuehl, Jr., E. H. Flynn, & K. Folkers. J. Am. Chem. Soc. 68: 2405. 1946.

[11] Fried, J., & O. Wintersteiner. Abstracts of the Chicago Meeting, A.C.S., Sept., 1946: 15 B.

[12] Hooper, I. R., L. H. Klem, W. J. Polglase, & M. L. Wolfram. J. Am. Chem. Soc. 68: 2120. 1946.

with hydroxylamine and with semicarbazide has been shown to occur.[13] Reduction of streptomycin gives rise to dihydrostreptomycin,[14] a biologically active substance, while oxidation yields biologically inactive streptomycinic acid.[11] These reactions apparently involve only the aldehydic group. A diagrammatic representation of the structure of streptomycin is shown in FIGURE 1.

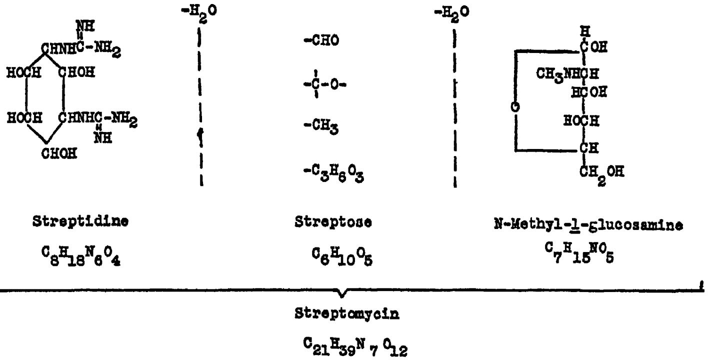


FIGURE 1.

Streptomycin trihydrochloride is quite soluble in both water and methanol, and slightly soluble in ethanol. The sulfate is quite soluble in water but only slightly soluble in methanol. Aqueous solutions of pure streptomycin trihydrochloride are slightly acidic. The activity of streptomycin trihydrochloride is about 800 units per mg., although the values obtained depend to a certain extent on the methods of assay.

Fried and Wintersteiner[2] first reported the preparation of crystalline streptomycin reineckate from concentrates of streptomycin purified by steps including chromatography. From their brief description, it appears that streptomycin picrate was employed for the chromatographic purification. The reineckate assayed about 400 units per mg., and streptomycin sulfate prepared from the pure reineckate assayed about 850 units per mg. Shortly afterwards, there appeared a paper by Kuehl, Peck, Walti, and Folkers[3] which described crystalline salts of streptomycin, including the helianthate, *p*-(2-hydroxy-1-naphthylazo)-benzenesulfonate, and sulfate, which were also prepared from streptomycin concentrates obtained by chromatographic procedures.

Carter and his associates[5] first described in detail a chromatographic purification procedure for the preparation of highly active concentrates of streptomycin. This method employed alumina as adsorbent and 80 per cent methanol as developing solvent. The streptomycin concentrate used in the chromatograms had an activity of 150-300 units per mg. It was found that alkaline alumina removed streptomycin from neutral aqueous solution, but elution with aqueous acid was slow and incomplete. Acid-washing alumina did not remove streptomycin from neutral aqueous solution, but did so from aqueous methanol. The chromatographic method used was developed on the basis of this information. It is to be noted that solutions of the concentrates of streptomycin hydrochloride were partially neutralized before introduction to the columns.

---

[13] Brink, N. G., F. A. Kuehl, Jr., & K. Folkers. Science 102: 506. 1945.

[14] Peck, R. L., C. E. Hoffhine, Jr., & K. Folkers. J. Am. Chem. Soc. 68: 1390. 1946.

Crude streptomycin hydrochloride in 70-80 per cent methanol at a pH of about 6.3 was percolated over a sulfuric acid-washed alumina column (pH 5-6). An inactive filtrate fraction giving a positive Sakaguchi test for guanido groups was followed by a fraction giving a negative Sakaguchi test. The subsequent fractions gave positive and strong such tests. The rapid rise of this test to a peak paralleled the increase in activity of the fractions. A small amount of activity remained on the columns. This could be removed by lowering the methanol content of the developing solvent. The material so eluted contained sulfate ion but no chloride ion. The evidence indicated that chloride ion was replaced by sulfate ion on the column, an exchange which rendered the product (streptomycin sulfate) less soluble in the developing solvent (70-80 per cent methanol), and thereby caused it to require the use of water for elution. Streptomycin sulfate is known to be much less soluble than the hydrochloride in methanol. The most active fractions from the column were concentrated to remove methanol, and lyophilized to give amorphous powders ranging in activity from 600-900 units* per mg. Less active fractions could be purified by repetition of the chromatographic procedures.

The alumina used (Merck or Harshaw) had been acidified to pH 6 with 50 per cent sulfuric acid and back washed with distilled water until the washings were sulfate-free. The pH of the final wash was 6.0. The alumina was put into the column as a slurry in water, then thoroughly washed with 80 per cent methanol. Columns, 3.2 cm. in diameter, containing 480 cc. of packed alumina were used for 8-10 gram lots of streptomycin. The weight ratio of adsorbent to solute is estimated to be about 25-50:1.

The crude streptomycin hydrochloride (8-10 g.) was dissolved in 90 cc. of 80 per cent methanol and brought to pH 6.3 with 2.0 *N* lithium hydroxide in 80 per cent methanol. The slight precipitate was removed by filtration, and the filtrate was carefully poured onto the prepared column. The chromatogram was developed with 80 per cent methanol. The solution leaving the column was collected in 100-cc. fractions. Sakaguchi and chloride tests were made to check the location of the fractions. When the Sakaguchi test began to decrease, the 80 per cent methanol was replaced with water as developing solvent. This accounts for the observation of secondary peaks of activity (see FIGURE 2). Various fractions were combined for preparation of solid products. TABLE 1 gives

---

* In such an instance, it is believed that comparative assays with pure streptomycin trihydrochloride would show the latter to be more active than 900 units per mg.

TABLE 1*

SOLID FRACTIONS FROM ALUMINA COLUMN

Fraction No.	Weight of solid g.	Activity of solid units per mg.	Nitrogen content per cent
7-11	0.45	600	11.40
12-17	1.00	900	13.47
18-23	0.65	620	13.83
24-26	0.43	520	

* From Carter, H. E., R. K. Clark, Jr., S. R. Dickman, Y. H. Loo, P. S. Skell & W. A. Strong. J. Biol. Chem. 160: 337. 1945.

data on these fractions. Based on the activity of pure streptomycin hydrochloride,[2, 3] the best fractions from such columns were probably at least 90 per cent pure streptomycin hydrochloride. TABLE 2 shows the

TABLE 2*

FRACTIONATION OF CRUDE STREPTOMYCIN ON ALUMINA COLUMN

Fraction No.	Volume cc.	Sakaguchi	Chloride	Activity units per cc.	Total units
2	800	—	+	0	
3	100	—	+	0	
4	100	+	+	0	
5	100	+	+	0	
6	100	+	+	0	
7	100	—	—	16	1,600
8	100	+	+	170	17,000
9	100	+	+	400	40,000
10	100	+	+	570	57,000
11	100	++	+	800	80,000
12	100	+++	+	1140	114,000
13	100	+++	+	1140	114,000
14	100	+++	+	1140	114,000
15	100	+++	+	1340	134,000
16	100	+++	+	1340	134,000
17	100	+++	+	1140	114,000
18	100	++	+	960	96,000
19	100	++	+	920	92,000
20	100	++	+	570	57,000
21	100	++	+	720	72,000
22	100	++	+	1060	106,000
23	100	+	+	570	57,000
24	200	+	+	800	160,000
25	200	+	—	380	76,000
26	200	+	—	400	80,000
Total units recovered					1,715,000

* From Carter, H. E., R. K. Clark, Jr., S. R. Dickman, Y. H. Loo, P. S. Skell & W. A. Strong. J. Biol. Chem. 160: 337. 1945.

fractionation data obtained in the chromatogram described above. The streptomycin hydrochloride put onto this column weighed 9.9 g., had an activity of 221 units per mg., and contained a total of 2.19 million units.

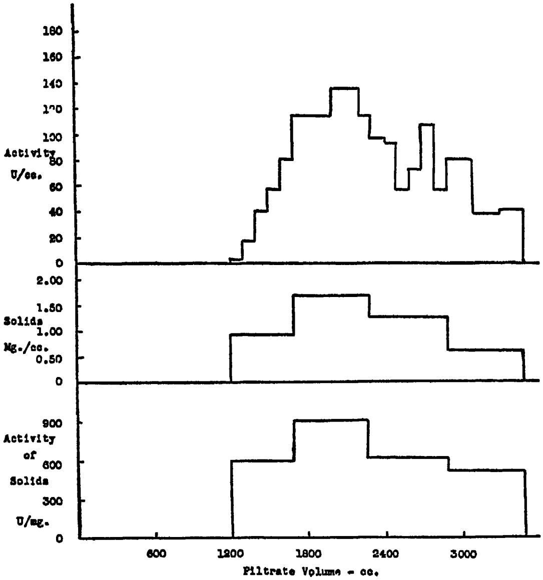


FIGURE 2. Chromatogram of streptomycin hydrochloride, 221 units/mg. Solvent: 80% methanol; adsorbent: alumina. (Data from Carter et al.[5])

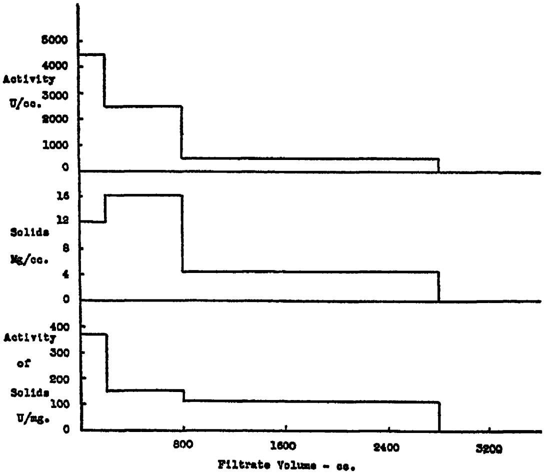


FIGURE 3. Chromatogram of streptomycin hydrochloride, 72 units/mg. Solvent: methanol; adsorbent: alumina. (Data from Kuehl et al.[15])

FIGURE 2 gives a graphic study of this chromatogram. The solid content of the filtrate is quite low, reaching a maximum of only about 1.7 mg. per cc. (fractions 12-17). The activity maximum in the filtrate was about 1340 units per cc. The second maximum of activity in FIGURE 2 is the result of change from 80 per cent methanol to water after fraction 20 had been taken. It will be observed, from the graphs, that the most active fractions were obtained at the point of maxima of both activity and solids concentration. Lack of data on solids concentration makes it hard to judge the degree of coincidence of these maxima. It is to be noted that considerable solvent was collected before much solid appeared in the filtrate. Lacking other data, this indicates a moderately slow passage of the solute through the column, relative to solvent flow. Since streptomycin is a base, the lower the pH of its solutions the more ionized it will be and, consequently, the less strongly adsorbed. The slow passage of the active material through this column, therefore, may be ascribed in part to its partial neutralization. Slow passage through the column should favor effective purification. Approximately four-fold purification was achieved in this experiment, and the best fractions were very close to purity, in so far as they may be judged by bioassay data.

Kuehl and associates[15] recently published details of their chromatographic purification procedures. They used absolute methanol solutions of streptomycin hydrochloride concentrates without preliminary neutralization, and absolute methanol as developing solvent. Darco G-60 carbon and sulfuric acid-washed alumina were employed as adsorbents. Filter paper pulp was mixed with the Darco G-60 for bulking purposes.[16] The filtrate fractions were collected and mixed with 5-10 volumes of acetone or ether to effect precipitation of the streptomycin hydrochloride. The products were collected by centrifugation, washed, and dried *in vacuo*. TABLE 3 gives some examples of the data obtained.

In FIGURES 3, 4, and 5 graphic studies of these results are shown. The maximum solids concentration in the filtrate fractions from these chromatograms is in the range of 7-16 mg. per cc., roughly 5-10 times as high as in the case of the column described by Carter *et al.*[5] The maximum activity concentration is in the range of 3000-8500 units per cc., again considerably higher than in the previously described case; also, this maximum occurs before the attainment of the maximum of solids. Owing, probably, to the lower pH used in these chromatograms, the active material passes down the column faster than in the preceding experiment.

In FIGURE 3, the activity maximum appears before the maximum of solids concentration, indicating a significant separation of streptomycin from one or more impurities. Separation of two zones is approximated. The concentrate employed in this column contained roughly only ten per

---

[15] Kuehl, F. A., Jr., R. L. Peck, C. E. Hoffhine, Jr., R. P. Graber, & K. Folkers. J. Am. Chem. Soc. 68: 1460. 1946.

[16] Wachtel, J. & H. G. Cassidy. Science 95: 233. 1942.

TABLE 3

CHROMATOGRAPHIC PURIFICATION OF STREPTOMYCIN HYDROCHLORIDE

Solute		Adsorbent			Products		
Wt. g.	Activity units/mg.	Wt. g.	Type	Eluate vol. cc.	Wt. of fraction g.	Activity units/mg.	Per cent of activity recovered
45	72	800	Alumina	200	2.42	372	28
				600	9.67	154	46
				2000	8.67	113	30
6.0	225	160	Alumina	250	0.698	500	23
				150	1.031	490	34
				250	0.729	290	14
				350	0.413	160	4
2.75	350	160	Darco G-60 filter paper pulp 10:1 mixture	25	0.202	780	16
				25	0.355	600	21
				50	0.820	400	33
				50	0.367	220	8
				250	0.208	210	4

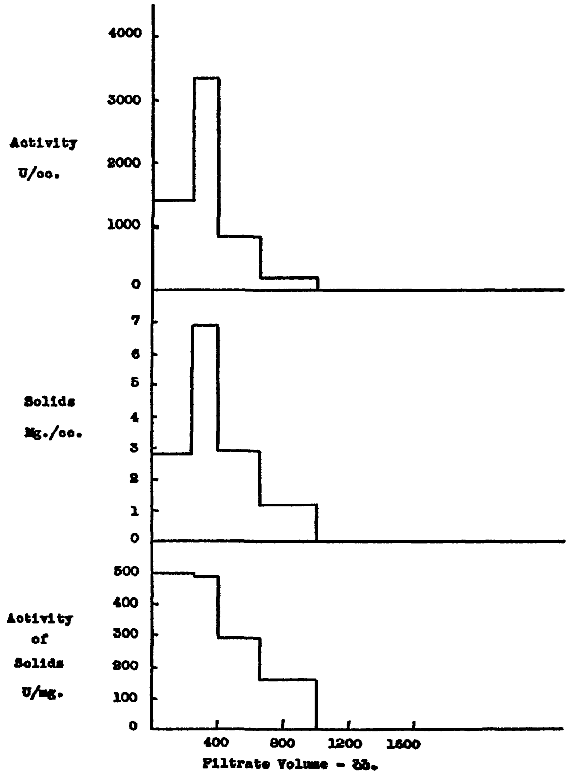


FIGURE 4. Chromatogram of streptomycin hydrochloride, 225 units/mg. Solvent: methanol; adsorbent: alumina. (Data from Kuehl et al.[15])

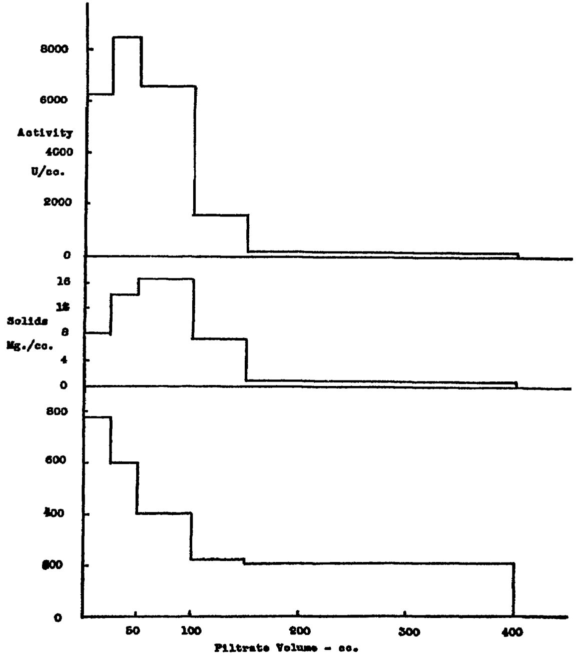


FIGURE 5. Chromatogram of streptomycin hydrochloride, 350 units/mg. Solvent: methanol; adsorbent: Darco G-60. (Data from Kuehl et al.[15])

cent of streptomycin, and the streptomycin appeared to be about the most rapidly moving component of the mixture. It is pointed out that, if there are salts present, *e.g.*, sodium chloride, they will move nearly as fast as the solvent, and faster than streptomycin. Their rate relative to solvent in this system[17] approaches unity; $R \cong 1$. FIGURE 4, a chromatogram of a purer concentrate on alumina, shows no definite separation of maxima. FIGURE 5, however, a chromatogram employing Darco G-60 carbon, again shows this separation. It is possible that the carbon effects greater separation of the components of streptomycin concentrates than does alumina. Its use on a large scale is, however, beset with certain practical difficulties.

Vander Brook and associates[7] have recently described a chromatographic procedure similar to that of Carter *et al.*[5] Their modification lies in the use of hydrochloric acid-washed alumina which eliminates the streptomycin sulfate zone noted[5] in the chromatogram on sulfuric acid-washed alumina. The procedure employed was the following. Five kilograms of Harshaw alumina (No. 2-350, catalyst) were adjusted to pH 4.7 with hydrochloric acid and thoroughly washed with distilled water. The acid-adjusted aluminum oxide was then suspended in 80 per cent methanol and poured, as a slurry, into a glass pipe, 4 inches in diameter, producing, after settling, a column 25 inches high.

A solution of 64 million units of streptomycin hydrochloride (160 g., 400 units per mg.) in water was brought to pH 5.8 in a final volume of 350 cc. The aqueous solution was diluted with 4 volumes of methanol before passage over the column. The weight ratio of adsorbent to solute was about 30:1. Development was carried out with 80 per cent methanol. The filtrate fractions were concentrated *in vacuo* to remove the alcohol, then dried from the frozen state. The results are shown in TABLE 4.

TABLE 4*

CHROMATOGRAPHIC SEPARATION OF 64 MILLION UNITS OF STREPTOMYCIN HYDROCHLORIDE (400 units per mg.) OVER ALUMINA COLUMN

Fraction No. (300 cc. each)	Units 10^6	Solids g.	Purity units per mg.	Recovery per cent
10	4.3	5.12	840	6.7
11	20.0	20.17	995	31.2
12	17.4	19.85	879	27.2
13	9.6	16.00	600	15.0
14	6.7	14.10	475	10.5

* From Vander Brook, M. J., A. N. Wick, W. H. De Vries, R. Harris, & G. F. Cartland. J. Biol. Chem. 165: 463. 1946.

---

[17] Le Rosen, A. L. J. Am. Chem. Soc. 64: 1905. 1942.

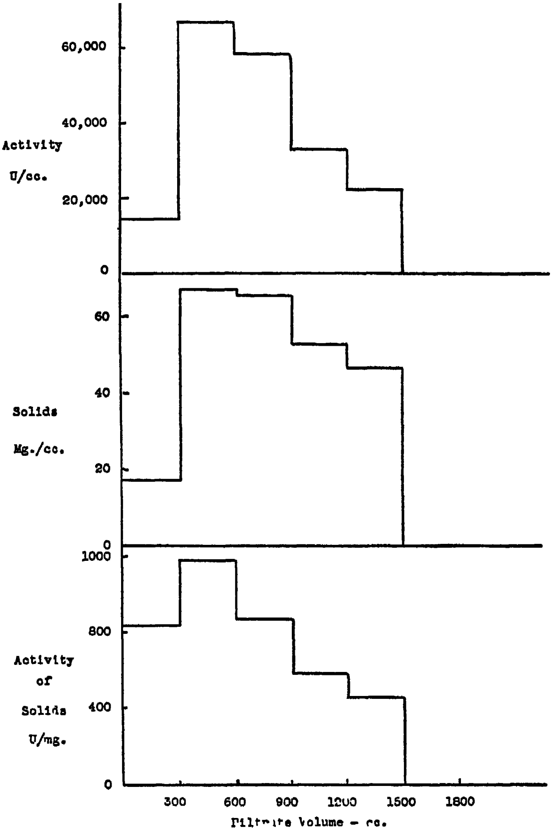


Figure 6. Chromatogram of streptomycin hydrochloride, 400 units/mg. Solvent: 80% methanol; adsorbent: alumina. (Data from Vander Brook et al.[7])

A graphic study of this column is shown in FIGURE 6. A very high concentration of material was used for this chromatogram. The solids in the filtrate attained a maximum of nearly 70 mg. per cc., and the filtrate activity reached a maximum value of about 67,000 units per cc. In this column, as in those previously described, the activity maximum appears to be ahead of the solids maximum. However, probably because of the high concentration of solute and its fairly high activity, this is not so obvious as in some of the other examples. The passage of active material was fairly slow, as judged by the fact that significant amounts of active material first appeared in fraction 10. The pH of the solution put onto this column was close to 6.0; hence, slow passage and very effective purification would be anticipated. In this chromatogram, fractions of streptomycin which were very nearly pure were obtained. This example is a good illustration of the preparative application of the chromatographic purification of streptomycin.

The chromatographic purification of streptothricin concentrates has been described by Peck *et al.*[18] The results obtained are of interest for comparison with the results of the chromatographic purification of streptomycin on alumina and Darco G-60. The procedures used were essentially identical with those described by Kuehl *et al.*[12] and were, in fact, the basis for the latter method. TABLE 5 shows the data obtained, and FIGURES 7 and 8 give graphic analyses of the data.

TABLE 5

CHROMATOGRAPHIC PURIFICATION OF STREPTOTHRICIN HYDROCHLORIDE

Solute		Adsorbent			Products		
Wt. g.	Activity units/mg.	Wt. g.	Type	Eluate vol. cc.	Wt. of fraction g.	Activity units/mg.	Per cent of activity recovered
10.0	49	160	Alumina	148	2.30	20	4
				30	0.76	64	10
				402	2.48	112*	57
				452	0.55	81*	9
2.76	120	77	Darco G-60 filter paper pulp (42:35)	70	0.959	40*	13
				42	0.570	131*	23
				43	0.381	186*	21
				152	0.431	233*	28

* Averaged values of composite fractions.

The maxima of solids content (14 and 25 mg. per cc., respectively, for Darco G-60 and alumina) and the maxima of activity (about 1800 and 7000 units per cc., respectively) are in about the same range as those described for streptomycin by Kuehl *et al.*[15] The concentrates used for the chromatograms illustrated in FIGURES 7 and 8 were of rather low activity. Comparison with FIGURES 3 and 5, for alumina and Darco G-60 carbon, clearly shows that, whereas the activity maximum appears before the solids maximum in the case of streptomycin, the opposite is true with streptothricin. The evidence indicates that the impurities in streptothricin move at a faster rate than the active substance. Since the pH of the concentrate solutions of both antibiotics were of about the same magnitude, streptothricin appears to be more strongly adsorbed than streptomycin.

Streptomycin and streptothricin are produced under very similar conditions by related microorganisms. It is reasonable to assume that, of the impurities present in crude concentrates, a considerable amount would

---

[18] Peck, R. L., A. Walti, R. P. Graber, E. H. Flynn, C. E. Hoffhine, Jr., V. Allfrey, & K. Folkers. J. Am. Chem. Soc. 68: 772. 1946.

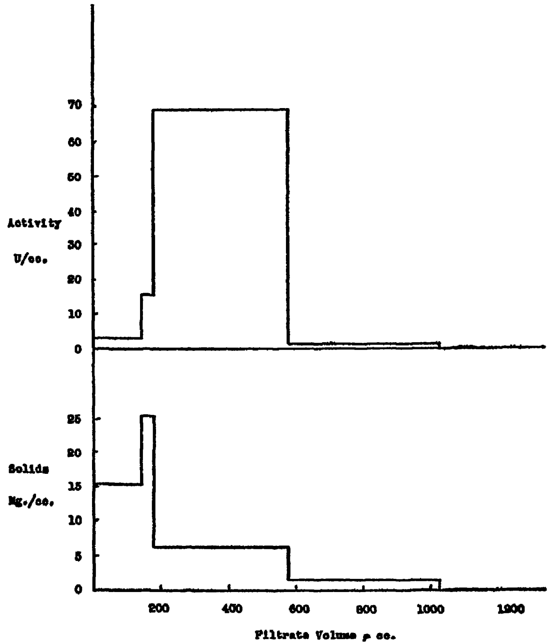


FIGURE 7. Chromatogram of streptothricin hydrochloride, 49 units/mg. Solvent: methanol; adsorbent: alumina. (Data from Peck et al.[18])

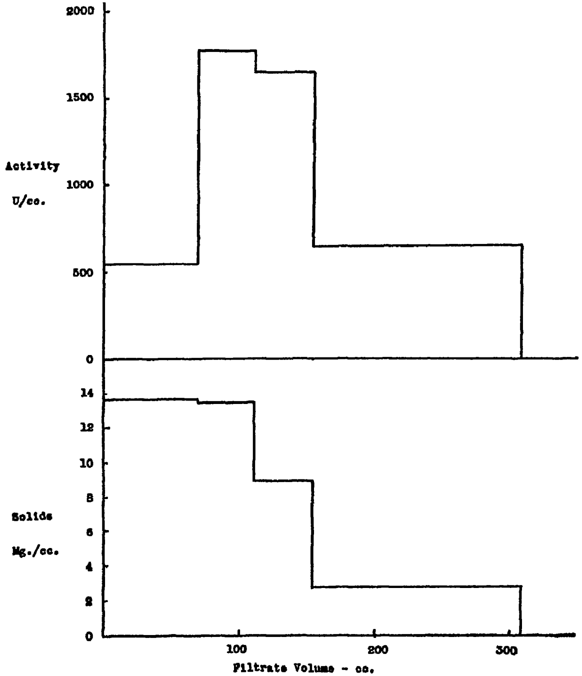


FIGURE 8. Chromatogram of streptothricin hydrochloride, 120 units/mg. Solvent: methanol; adsorbent: Darco G-60. (Data from Peck et al.[18])

be the same for equivalent concentrates of the two antibiotics. On the basis of this assumption, and from the data compared above, it would be anticipated that a mixture of streptomycin and streptothricin could be separated chromatographically by the systems described above and that the streptomycin would appear first in the filtrate fractions.

When it was found that the most active fractions obtained by chromatography of streptomycin hydrochloride in methanol on alumina or Darco G-60 were in the first filtrate fractions collected, it was apparent that streptomycin was about the least strongly adsorbed component of crude concentrates. The conclusion was drawn[15] that, by adding portions of alumina to a methanolic solution of streptomycin hydrochloride, a substantial amount of the impurities present would be removed, thus effecting an increase in the activity of the streptomycin remaining in solution. This proved to be true, although the efficiency of this purification method was somewhat less than that of the chromatographic procedure.

Mueller[19] recently described a moderately large-scale application of this method, employing streptomycin phosphate-hydrochloride in absolute methanol. Washed Harshaw alumina (passing 80 mesh) was used as adsorbent. Streptomycin preparations of 200-300 units per mg. could be increased in activity to any desired degree up to a practical limit of 650 units per mg. Adsorption isotherms for two crude streptomycin salts were determined by Mueller[19] for various conditions. Curves showing activities, together with total recovered solids and activities, permit the yield and activity to be predicted for the product of any operation. The method is useful where the highest activities are not desired.

SUMMARY

The chromatographic purification of streptomycin has been conducted with acid-washed alumina and with Darco G-60 carbon as adsorbents and with methanol or 80 per cent methanol as developing solvents. Concentrations employed varied over a wide range. An example of fairly large-scale chromatography of a concentrate has been described. Streptomycin appears to pass down chromatographic columns at a rate faster than that of most of the other components of the crude concentrates used, that is, it was least strongly adsorbed. The rate does, however, appear to depend somewhat on the pH of the solutions employed, and is slower at the higher pH values. Batch treatment of methanol solutions of streptomycin with alumina removed most of the impurities and left the active material in solution in a form readily isolated.

---

[19] Mueller, G. P. Abstracts of the Chicago Meeting, A. C. S., Sept., 1946: 14B.

PARTITION CHROMATOGRAPHY

By A. J. P. MARTIN

Boots Pure Drug Company Limited, Nottingham, England

Synge,[1] in 1938, was preparing the acetyl derivatives of the amino acids and showed that there were considerable differences between the partition coefficients of many of them between $CHCl_3$ and water. He felt that these differences could be made the basis of a scheme of separation of the more fatty mono-aminomonocarboxylic acids. Martin, some time previously, had constructed a counter-current extraction machine for vitamin purification. This machine, though at its best it had an efficiency of some 200 theoretical plates, was unsuitable for the chloroform-water system needed for the acetyl amino acids. Martin and Synge,[2] therefore, constructed a different type of machine with an efficiency of less than 40 theoretical plates, which was used for the separation of certain of the amino acids in a hydrolysate of wool. However, the quantity of 40 theoretical plates was inadequate to obtain clean-cut fractions. The mechanical and running difficulties, at the same time, precluded the use of a machine of much higher efficiency, and other methods were tried.

The essential trouble with the machines was that, if the size of the drops of the dispersed phase was small enough to permit equilibrium to be attained in a reasonable time, then it was difficult to obtain the required relative movement of the two phases, *i.e.*, entrainment occurred or the settling time was excessive.

An attempt was made to overcome this trouble by using a column packed with parallel fibers of wool and cotton, with the idea that the cotton would wet with the aqueous phase and the wool with the organic phase, and that many parallel counter-current streams could be maintained in a small tube. Not surprisingly, perhaps, this scheme did not work and large areas became the domain of one phase only.

The idea of obtaining flow in both directions was therefore abandoned, and an attempt was made to hold one phase stationary while the other was allowed to flow past it.

Silica gel was chosen to hold water stationary while chloroform flowed past it. The silica would hold about half its weight of water without apparently becoming wet, and chloroform did not tend to displace the water. It was possible, therefore, to make up a "partition chromatogram" with silica gel containing water and, using chloroform as the mobile phase, to effect a separation between various acetyl amino acids, making use of their difference in partition coefficients between chloroform and water. It was, of course, necessary that the silica itself should not adsorb too much of the acetyl amino acids and we were fortunate in that, for the first experiment, a commercial chloroform was used which contained appreciable amounts of ethyl alcohol. In a second experiment using purified chloroform, the acetyl amino acids were all held at the top of the column, and it was found, later, that some relatively polar compound had to be present in the solvent to act as an eluting agent, to prevent the silica, as far as possible, from acting as an adsorbent.

There were, of course, several materials which could have been used to hold water in this way. Silica, however, had one particular advantage. The acetyl amino acids being quite strong acids, it was feasible to reveal what part of the column they occupied by including, in the water in the silica, an indicator such as methyl orange. The column then turned to red (from yellow or orange) in the regions where the acetyl amino acids were present. It was not possible to make this indicator-color change satisfactory in any other water-holding substance, the range of the indicator usually being shifted to a lower pH. The question of suitable indicators will be discussed below.

Since the problem of the separation of the acetyl amino acids had been approached from the angle of the counter-current extraction machines, it was natural to try to compare the respective efficiencies of the machines, as well as the partition chromatograms. A theory was therefore developed[3] expressing the efficiency of the chromatogram in terms of the number of theoretical plates of an extraction apparatus to which it was equivalent. For an average column, it was found to be of the order of thousands, this high efficiency being, of course, directly referable to the smallness of the distance that the solute had to diffuse to get into and out of either phase. It is, therefore, improbable for analytical purposes that any extraction machine can compete with the partition chromatogram, where the latter is suitable.

To return to indicators. One of the desirable characteristics of an indicator is insolubility in the mobile phase. In this respect, methyl orange is unsuitable with solvents containing high proportions of alcohol. For use with such solvents, Gordon, Martin, and Synge[4] introduced the use of the anthocyanins, peonin and pelargonin, while Liddell and Rydon[5] used 3:6-naphthalene-azo-N-phenylamine. These indicators, particularly the latter which is easily prepared, are satisfactory, being washed from the column slowly enough to permit the estimation of the required acetyl amino acids.

The indicator imposes, apart from questions of solubility, some limitation on the solvent systems that can be employed. With methyl orange, chloroform shifts the color towards yellow, and alcohols do so still more. Cyclohexane alone does not have much effect and, hence, larger proportions of alcohols have to be used with cyclohexane than with chloroform to obtain a usable color of the indicator. In general, alcohols of smaller molecular weight have relatively more effect on the color of the indicator and less effect on the partition coefficient of the acetyl amino acids than do alcohols of larger molecular weight.

From the theory given by Martin and Synge,[8] it is easy to show that if:

A is the area of cross section of the column,

A_S is the area of cross section of the non-mobile phase,

A_L is the area of cross section of the mobile phase,

A_I is the area of cross section of the inert solid ($A_S + A_L + A_I = A$),

α is the partition coefficient, *i.e.*, $\dfrac{\text{g. solute/l. of non mobile phase}}{\text{g. solute/l. of mobile phase}}$

$$R = \frac{\text{movement of position of maximum concentration of solute}}{\text{simultaneous movement of surface of developing phase in empty part of tube above chromatogram column}},$$

$$\text{then, } R = \frac{A}{A_L + \alpha\ A_S} \text{ or } \alpha = \frac{A}{RA_S} - \frac{A_L}{A_S}$$

The authors found[8] that, in the case of a particular column, the partition coefficient deduced from the rate of movement of the bands agreed very closely with that determined directly. This was taken to show that, under favorable conditions, adsorption could be practically eliminated, and that the water in the silica had the same solvent properties as normal water. In TABLE 1 are given the partition coefficients found for acetylproline and acetylphenylalanine, as well as those determined directly in the same solvent mixture. TABLE 2 shows the R values for a number of acetyl amino acids and solvents.

TABLE 1

Acetyl amino acid	R	α determined from R	α determined directly
Acetylproline	0.37	9.4	9.5
Acetylphenylalanine	1.07	1.4	1.3

TABLE 2

BAND RATES OF ACETAMINO ACIDS WITH VARIOUS SOLVENTS

Figures are values of *R*

Developing solvents	Butanol-chloroform			Propanol-cyclohexane		Ethyl acetate
	1%	3%	17%	5%	30%	
Acetylphenylalanine	0.5	0.9	Fast	0.3	Fast	Fast
Acetyl*nor*leucine	0.4	—	Fast	—	—	—
Acetylleucine Acetyl*iso*leucine	0.3	0.6	Fast	0.3	Fast	Fast
Acetyltryptophane	0.3	0.6	Fast	0.15	Fast	—
Acetyl*nor*valine Acetylmethionine	0.15	0.3	Fast	0.15	Fast	0.9
Acetylmethionine	0.15	0.3	Fast	0.09	0.7	0.8
Acetylproline	0.15	0.3	Fast	0.04	0.3	0.2
Acetaminobutyric acid	0.07	—	—	—	—	0.5
Acetylalanine	0.025	0.04	0.35	0.04	0.3	0.2
N-Acetyltyrosine	0.02	0.04	0.7	—	0.6	Fast
Acetylglycine	Slow	Slow	0.15	Slow	0.15	0.1
Acetylaspartic acid						
Acetylglutamic acid	Slow	Slow	0.1-0.2	—	—	—
N,N Diacetylcystine						
N,N Diacetyllysine						
N,Acetylhydroxyproline	Slow	Slow	0.07	Slow	0.1	0.06

The results given above have been shown to only one significant figure because, unless great care is taken in manufacture of the silica, moisture content, and in preparation of the column, the values are not consistent to better than this figure. The ratios are, however, much more consistent and, if several of the appropriate amino acids are present, there is never any difficulty in identifying a given zone; when only one zone is visible, it is desirable to confirm the identification by running chromatograms of the unknown mixed with an authentic sample of the substance that a preliminary run has indicated it to be. The question of the preparation of the silica will be taken up again below.

The procedure in analyzing a protein hydrolysate for the higher monoaminomonocarboxylic acids was as follows: The protein was hydrolyzed in a large excess of 6N HCl, boiling for 24 hours, and the HCl was largely removed by repeated distillation *in vacuo*, with additions of distilled water. The hydrolysate, corresponding to 25 mg. of protein, was then acetylated with five successive portions of Ac_2O and NaOH, the mixture being maintained ice-cold and alkaline throughout. It was then made strongly alkaline for 10 minutes at room temperature, to hydrolyze any O-acetylation that might have occurred. Finally, it was acidified with 10 N H_2SO_4 and extracted, in a separatory funnel, with 17% BuOH-$CHCl_3$ mixture. (If acids slower moving than alanine are required—and this will be only if but one or two such acids are present—

complete extraction, which cannot conveniently be obtained in a separatory funnel, can be obtained by taking up the acetylation mixture in dry silica, making it into a column, and extracting with the BuOH-$CHCl_3$ mixture. A 1-cm. layer of water-saturated silica, introduced by Tristram,[6] is required below the main column to prevent any mineral acid from being extracted.)

The BuOH-$CHCl_3$ solution of acetyl amino acids was evaporated down and taken up in one or two ml. of 1% BuOH-$CHCl_3$. It was now ready to be applied to the column, which was prepared as follows: To 3 grams of dry silica in a 100-ml. beaker were added 1.5 ml. of a saturated solution of methyl orange, and made homogenous through light grinding with a test tube. 1% BuOH-$CHCl_3$ was added until a cream was produced, and this was then poured into the chromatogram tube. When the column had settled and drained, no entry of air into the column occurred, provided that the lower end of the tube was not allowed to become quite dry. Acetyl amino acid solution was then run carefully onto the column and allowed to soak in, and followed by three rinsings with similar volumes. The top of the column had now changed in color from orange-pink to bright red. The column was then developed with 1% BuOH-$CHCl_3$ mixture, and zones of acetylphenylalanine, acetylleucine + acetyl*iso*leucine, acetylproline + acetylvaline + acetylmethionine, and acetylalanine could be distinguished, with the other acetyl amino acids left at the top of the column. The first three of these zones were run out and collected separately. The point at which a cut should be made, could be determined by estimating, while zones on either side of the proposed cut were still visible, the amount of solvent standing above the column which should run in, so that the cut be made at the desired *R* value. With 1% BuOH-$CHCl_3$, an excessive volume would have been required to run out the acetylalanine zone, so a change was made to 17% BuOH-$CHCl_3$. Acetyltyrosine then overtook the acetylalanine, and both zones were collected separately.

After evaporation of the solvent, the fraction containing acetylproline, acetylmethionine, and acetylvaline was run on a column using 5% PrOH-C_6H_{12}. Acetylvaline and acetylmethionine were then collected separately, and acetylproline was run out by changing the solvent to 30% PrOH-C_6H_{12}. It was also found advisable to put the other zones from the $CHCl_3$ column through the cyclohexane column, which separated out some non-acetyl amino acid impurities. Finally, after evaporation of the solvents, the individual fractions were titrated against $\frac{N}{100}$ $Ba(OH)_2$ solution, using bromophenol blue as indicator.

In this way, quantitive analyses of 25 mg. of protein hydrolysate for phenylalanine, leucine and *iso*leucine, valine, methionine, proline, tyrosine, and alanine were obtained. The order in which the various solvents were used may be changed in certain cases (see reference 6).

It was found, from analyses of control mixtures of amino acids treated exactly as the protein, that results within 5 per cent of true values were obtained. It was thus considered desirable always to compare results obtained with a given protein with those obtained from a control mixture made up to simulate as closely as was possible the protein under investigation. Correction factors, derived from the results of the analysis of the control mixture, were then applied to the results of the protein analysis, and some further improvement of accuracy probably resulted.

All the sources of error have not been discovered. It was found that results obtained from different acetylations differed more than those obtained from the same acetylation, and that much better results could be obtained from mixtures of pure acetyl amino acids. Apart from failure to attain complete acetylation, there is a possibility of peptide synthesis as a consequence of the formation of a mixed anhydride, by reaction between acetic anhydride and the carboxyl of the amino acid, or acetyl amino acid.

There is also evidence of the presence, in the acetylated hydrolysate, of acids containing less nitrogen than the acetyl amino acids and, presumably, derived from decomposition products of the protein.

Tristram[6] has also found great variation in the recovery of methionine, due, probably, to oxidation of the sulphur. A discussion of his results in an extensive series of analyses is given in the same paper.

However, probably the biggest source of error lies in the silica. It has not been found possible to devise a method of consistently preparing satisfactory silica, though, recently, new methods have been proposed of which the author has no experience. The silica may fail either in that the color of the indicator is such that the acetyl amino acid zones are unobservable, or in its absorbing the acetyl amino acid with the result that a zone of one acetyl amino acid tails into the succeeding zone. Tests (Gordon, Martin and Synge,[4] Tristram[6]) involving separations of test mixtures have been proposed for excluding unsatisfactory batches. Different batches of water glass used for the preparations of the silica require different methods, and further work on this subject is needed. Isherwood[7] believes that many of the troubles are due to the presence of iron and, in his method, takes particular pains to remove this.

The partition chromatogram on silica has been used for the estimation of a number of other acids Hanby and Rydon[8] have converted glutamic acid to pyrrolidone carboxylic acid by autoclaving at 100° for 4 hours in a solution brought to pH 2 with acetic acid. The pyrrolidone carboxylic acid can then be run with 17% BuOH-$CHCl_3$ ($R = 0.3$) and overall recoveries of more than 95 per cent are obtainable.

Elsden[9] has used the method for the smaller fatty acids, taking the method for 1 or 5% BuOH-$CHCl_3$ and bromocresol green as an indicator. The acids are titrated using a CO_2-stirred emulsion. TABLE 3 shows the recoveries he has obtained.

TABLE 3

	Total acid (ml)	Butyric acid (ml)	Propionic acid (ml)	Acetic acid (ml)
Theoretical	28.93	9.31	9.58	10.04
Observed	28.86 ± 0.35	9.4 ± 0.18	9.52 ± 0.26	9.93 ± 0.31

Patterson[10] has also used the silica partition for the separation of fatty acids, with satisfactory results. Sanger[11] has used silica columns with a variety of solvents for the separation of dinitro phenyl derivatives of amino acids. There is reason to believe, however, that, in his case, adsorption on the silica was playing a more important role than partition between the solvents and water. When silica suitable for use with the acetyl amino acids is employed, the *R* values are all much higher than those quoted by Sanger, and no useful separations are obtained. This is inexplicable if partition is the dominating phenomenon. Isherwood[7] has used the partition chromatogram on silica for the separation of a number of hydroxy acids from fruits. He has found it desirable to work with dilute sulphuric acid on the silica and, hence, cannot use an indicator in the column. He has found, however, that it is possible to run the effluent from the column down a short capillary tube with a slow stream of dilute alkali containing an indicator. The color of the indicator as it emerges from the capillary shows whether acid is running from the column. This method is potentially of very wide application, and its use with the acetyl amino acid separations might well avoid several of the snags associated with the use of silica. The accuracy of Isherwood's analyses for hydroxy amino acids is comparable with those of the acetyl amino acids.

Partition chromatograms have also, of course, been widely used for the separation and purification of the various penicillins. A buffer is used in the silica in place of water, and various solvents and buffers have been used. Most of the literature on this is not as yet published (see reference 12).

PARTITION CHROMATOGRAPHY ON PAPER AND STARCH

Gordon, Martin, and Synge[13] were not able to separate the slower moving acetyl amino acids on silica and turned to the problem of the separation of the amino acids on partition columns. Following Dakin's work, and also that of England and Cohn, the first solvent tried was, of course, butyl alcohol. However, the silica was found to adsorb the amino acids too strongly, particularly those of higher molecular weight.

Of course, no indicator could be used for the amino acids and, hence, one of the chief reasons for preferring silica to other substances for holding the water was no longer valid.

The first choice among other possible materials was cellulose. Using a semicircle of filter paper as in the well-known technique of capillary analysis[14], but ensuring that the cellulose was saturated with water by carrying out the experiment in a closed box whose atmosphere was kept saturated with water vapor, it was found readily possible to separate various mixtures of amino acids, and the separation could be revealed after drying the paper by spraying with a solution of 0.1 per cent ninhydrin in BuOH and heating. The amino acids were then shown up as colored rings. Later, it was found more convenient to use strips of paper hanging from a trough containing the solvent saturated with water, the whole being enclosed in a vessel, the air within which was saturated with respect to both solvent and water. This work was continued by Consden, Gordon, and Martin.[15]

It was found, in the case of BuOH as solvent, that the rate of movement of the amino acid could be predicted from the amounts of water and solvent on the paper, and from the partition coefficient of the amino acid.

The relation of the partition coefficient to the rate of movement of the band may be calculated by the method of Martin and Synge.[3]

A =cross-sectional area of paper+water+solvent,

A_L =cross-sectional area of solvent phase,

A_S =cross-sectional area of water phase,

α =partition coefficient, $= \dfrac{\text{concentration in water phase}}{\text{concentration in solvent phase}}$,

$$R = \frac{A}{A_L + \alpha A_S}$$

However, R is not conveniently measurable in paper chromatograms, so that a new symbol, R_F was introduced:

$$R_F = \frac{\text{movement of band}}{\text{movement of advancing front of liquid}} = \frac{RA_L}{A} = \frac{A_L}{A_L + \alpha A_S},$$

$$\text{or } \alpha = \frac{A_L}{R_F A_S} - \frac{A_L}{A_S} = \frac{A_L}{A_S}\left(\frac{1}{R_F} - 1\right).$$

A_L/A_S is equal to the ratio of the volumes of solvent and water phase in the chromatogram.

TABLE 4

PARTITION COEFFICIENTS CALCULATED FROM R_F VALUES

No. of run	1	2	3	4	Direct measurements (England & Cohn*)
Per cent water in paper	28.7	18.0	22.6	17.7	
A_L/A_S	3.25	4.56	3.70	2.93	
Amino acid	*Partition coefficients*				
Glycine	70.4	70.4	70.4	70.4	70.4
Alanine	35.9	39.9	43.7	36.6	42.3
Valine	12.2	14.1	14.8	12.5	13.8
*Nor*valine	8.7	10.8	10.5	9.2	9.5
Leucine	4.5	5.4	5.6	6.0	5.5
*Nor*leucine	3.5	4.2	4.4	4.6	3.2

* J. Am. Chem. Soc. 57: 634. 1935.

Assuming a given water content of the paper, A_L/A_S may be deduced from the ratio of weight of dry paper to that of the developed chromatogram.

The water content of the paper is apt to vary from experiment to experiment and is difficult to measure in the presence of *n*-butanol or another solvent. TABLE 4 shows the partition coefficient calculated from the R_F and A_L/A_S values for four separate runs under slightly different conditions. The water content has been assumed so that the partition coefficient for glycine is equal to that given by England and Cohn (1935). The last column gives the direct measurements of England and Cohn.

The water content of saturated cellulose according to the *International Critical Tables,* is 22 per cent on a dry-weight basis.

TABLE 5 shows a number of R_F values for various amino acids and the most commonly used solvents. Some additions have been made in some solvents, *e.g.,* cupron, HCN, and coal gas. These were to overcome a puzzling and unexpected difficulty. In many of the early chromatograms run with phenol, butyl alcohol, and benzylalcohol, a purple spot showing the amino acid was preceded by a pink "beard" of rather faster-running material, and, in general, it was found that the further the amino acid had run, the larger was the pink spot and the smaller the purple one. Thus, by the time leucine had run to the bottom of the sheet, it might be almost completely converted to a rather ill-defined pink streak. This was found to be due to the combination of the amino acid with copper which was in the paper, the copper complex having a higher R_F value than the corresponding amino acid. This trouble could be overcome by including in the atmosphere or solvent some substance that would form a more stable complex with the copper than the amino acid copper complex. It was found later[16] that, in the case of cysteic acid, zinc also formed a complex, though this was slower-running than the amino acid.

In TABLE 5, it will be noticed that the R_F values in collidine of an amino acid and the corresponding hydroxy amino acid are very similar, whereas the R_F values in phenol are considerably depressed by the addition of a hydroxyl group to the molecule. This indicates that little energy is required to transfer a hydroxyl group from water to collidine, but relatively much from water to phenol. A comparison of aromatic and aliphatic amino acids in butanol and benzyl alcohol shows similar differences. The effect of including NH_3 in the atmosphere on the R_F values of the basic and acidic amino acids also follows the expected pattern. It is hoped that, as further experience with different solutes and solvents is obtained and information about energies of association of various groups increases, a rational approach will make it possible to select the most effective solvent for a given separation. It may be that the partition chromatogram will, in this way, prove to be more predictable and even more versatile than the adsorption chromatogram.

It will be seen in TABLE 5 that the order of the various acids changes in different solvents. This suggested the possibility of further useful

TABLE 5

R_F VALUES OF AMINO ACID IN VARIOUS SOLVENTS ON WHATMAN NO. 1 PAPER, AT ROOM TEMPERATURE

Solvent	Phenol	Phenol	*s*-Collidine	*n*-Butanol	Benzyl alcohol	*o*-Cresol	*o*-Cresol	*m*-Cresol	*m*-Cresol	*p*-Cresol	*p*-Cresol	*Iso* butyric acid
Addition	0.1% NH_3* coal gas	Coal gas	—	Cupron	HCN	Cupron	Cupron 0.1% NH_3*	Cupron	Cupron 0.1% NH_3*	Cupron	Cupron 0.1% NH_3*	—
Glycine	0.41 RP	0.40	0.25 P	0.05	0.02	0.07	0.07	0.12 P	0.14	0.15 RP	0.19	0.36
Alanine	0.55 P	0.57	0.32 P	0.08	0.03	0.13	0.15	0.23 P	0.24	0.28	0.32	0.44
*Nor*valine	0.78 P	0.78	0.48 P	0.26	0.12	0.45	0.45	0.60 B	0.62	0.66 BP	0.66	0.71
Valine	0.76 P	0.78	0.45 P	0.20	0.11	0.41	0.38	0.52 P	0.54	0.59 BP	0.60	0.65
*Nor*leucine	0.87 P	0.85	0.60 P	0.47	0.27	0.72	0.67	0.80 B	0.78	0.82 BP	0.80	0.79
*Iso*leucine	0.87 P	0.82	0.54 P	0.37	0.18	0.58	0.55	0.70 B	0.70	0.73 BP	0.74	0.76
Leucine	0.86 P	0.84	0.58 P	0.38	0.21	0.61	0.62	0.73 B	0.73	0.73 BP	0.77	0.78
Phenylalanine	0.90 P	0.86	0.59 G	0.43	0.36	0.81	0.74	0.82 GB	0.82	0.82 P	0.83	0.80
Tyrosine‡	0.66 GP	0.59	0.64 G	0.28	0.14	0.24	0.23	0.35 P	0.35	0.39 B	0.38	0.58
Serine	0.35 P	0.33	0.28 G	0.05	0.01	0.05	0.05	0.08 BP	0.08	0.11 P	0.13	0.34
Threonine	0.47 P	0.50	0.32 P	0.07	0.02	0.11	0.08	0.14 B	0.14	0.19 P	0.21	0.43
Hydroxyproline	0.66 O	0.66	0.34 GY	0.07	0.04	0.27	0.23	0.32 GY	0.32	0.40 BG	0.42	0.42
Proline	0.89 Y	0.87	0.35 Y	0.12	0.12	0.69	0.66	0.73 Y	0.75	0.76 Y	0.78	0.57
Tryptophane	0.80 P	0.76	0.62 P	0.35	—	0.65	0.58	0.76	0.69	0.70 B	0.68	—
Histidine‡	0.70 RP	0.72	0.28 G	0.06	0.02	0.25	0.33	0.34 P	0.44	0.36 B	0.52	0.45
Arginine	0.85 P	0.67	0.16 P	0.03	0.01	0.02	0.44	0.07 P	0.76	0.12 BP	0.81	0.40
Ornithine‡	0.61 P	0.40	0.13 BG	0.01	0.00	0.01	0.21	0.03 P	0.47	0.04 P	0.59	0.24
Lysine‡	0.73 P	0.50	0.14 BG	0.01	0.00	0.01	0.20	0.04 P	0.54	0.07 BP	0.66	0.27
Aspartic acid**	0.12 BP	0.14	0.22 B	0.01	0.00	0.00	0.00	0.01 RP	0.01	0.02 BP	0.02	0.31
Glutamic acid	0.19 P	0.24	0.25 P	0.01	0.00	0.01	0.01	0.01 RP	0.02	0.03 BP	0.03	0.38
Lanthionine‡	0.21 RP	0.18	0.12 G	0.01	0.00	0.00	0.01	0.01 R	0.03	0.03 RP	0.05	0.21
Cystine‡	0.13 RP	0.13§	0.14 G	0.01	0.00	0.04	0.01	0.04 P	0.04	0.02 RP	0.02	0.25
Methionine	0.83 P	0.82	0.57 GP	0.26	0.17	0.58	0.49	0.64 B	0.63	0.71 BP	0.68	0.69

Color given by ninhydrin: B, Blue; G, Grey; O, Orange; P, Purple; R, Red; Y, Yellow.

* The % figure refers to the strength of NH_3 solution present in quantity in the tray.

‡ These acids were put on as the hydrochlorides and neutralized with NH_3 before development.

§ Cystine is decomposed by these additions. R_F value refers to the decomposition product.

** Color initially green in presence of phenol collidine. Only final color shown in table.

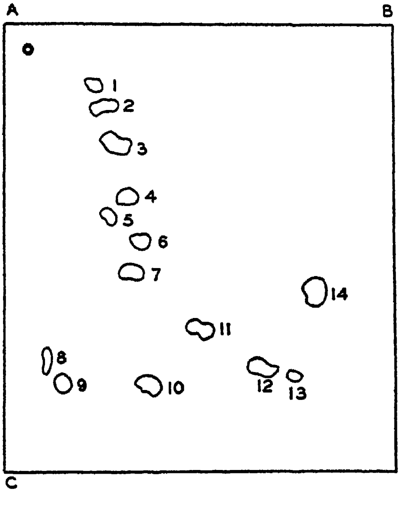
separation by two runs at right angles with different solvents, thus producing a two-dimensional chromatogram.

A sheet of filter paper (Whatman No. 4) 18 by 22″ is taken and near one corner a drop of protein hydrolysate is placed. The sheet is then hung from the trough and run with *s*-collidine. After a sufficient length of time (12 to 36 hours, depending on kind of paper and solvent), the paper is dried, turned through a right angle, and again run with phenol (in an atmosphere of coal gas with some NH_3 present). It is again dried, sprayed with ninhydrin solution and heated. Colored spots appear, showing the positions of the various amino acids. With the pair of solvents mentioned, cystine, cysteic acid, aspartic acid, glutamic acid, serine, glycine, threonine, alanine, hydroxy proline, proline, hydroxylysine, lysine, arginine, histidine, valine, and tyrosine appear as separate spots, while *iso*leucine, leucine, methionine, and phenylalanine and tryptophane fall together in an elongated spot. By using benzyl alcohol and butyl alcohol as the two solvents, those acids which are not separated by the collidine phenol chromatogram can be separated, though other amino acids will not then be resolved. The amount of protein hydrolysate required for one such two-dimensional chromatogram is about 300 micrograms. The amount of an individual amino acid that can be seen is 1 to 2 micrograms, and on a one-dimensional strip, half a microgram can, under favorable circumstances, be observed. One of these two-dimensional chromatograms is illustrated and interpreted in FIGURE 1.

In many instances, of course, the small quantity required for an analysis is a great advantage. Thus, it is possible to get much information from small amounts of valuable products, and to control fractionations by other methods, while using negligible aliquots. It is obvious that this method can be used in testing for other amino acids in allegedly pure amino acids.

The method is well adapted to any analysis for amino acids in which qualitative information is all that is required. Thus, Consden, Gordon, Martin, Rosenheim, and Synge[16] examined the samples of Thudichum's "glycoleucine" which had been isolated by him in 1880 from brain tissue and whose nature was uncertain, but which had come in many quarters to be regarded as *nor*leucine. Running comparative chromatograms of glycoleucine, *nor*leucine, leucine, and various mixtures of these, it was rapidly ascertained that glycoleucine was leucine. Other methods confirmed this and showed it to be *dl*-leucine (since it was from a baryta hydrolysate, this was, of course, only to be expected in the light of knowledge not available in 1880). A search was then made for *nor*leucine in hydrolyzed spinal cord, by the method of running the hydrolysate alone, *nor*leucine alone, and mixtures of the two, using benzyl alcohol as the solvent. It was shown that not more than .03 per cent of the nitrogen could be present as *nor*leucine, since when ½ microgram of *nor*leucine was added to 1.2 mg. of hydrolysate it could be clearly recognized, but none could be detected in the hydrolysate alone.

FIGURE 1. Two-dimensional chromatogram of a wool hydrolysate (180 μg.) on Whatman No. 1 sheet. Hydrolysate applied at circle. Run with collidine for 3 days in direction *AB*, then in direction *AC* with phenol for 27 hours in an atmosphere of coal gas and NH_3 (produced from a 0.3 per cent NH_3 solution). The filter employed in photographing renders the yellow proline spot scarcely visible. (From Biochem, J. 38, Plate 1, opp. p. 226. 1944.)



1 = Cystine
2 = Aspartic acid
3 = Glutamic acid
4 = Serine
5 = Glycine
6 = Threonine
7 = Alanine
8 = Lysine
9 = Arginine
10 = Proline
11 = Valine
12 = Leucine
Isoleucine
Methionine
13 = Phenylalanine
14 = Tyrosine

The paper partition chromatogram has been applied by Dent[17] to an extensive investigation of pathological urine, particularly in Fanconi's syndrome. The method has also been used by Edman[28] for the investigation of hypertension.

Partridge[18] has adapted the method to the investigation of the reducing carbohydrates. The same solvents that are useful for the amino acids are, surprisingly, useful also for the sugars. Instead of ninhydrin, Partridge sprays the paper with ammoniacal silver nitrate. The unreduced silver is then washed away, and a brown stain is left where the carbohydrate had been. Chargaff[19] has applied the method to the analysis of nucleotides, and Fink[20] to the separation of fatty acids after conversion to the corresponding hydroxamic acids.

Elsden and Synge suggested the use of potato starch for the separation of amino acids. In view of the development of this method by Moore and Stein, described in this publication, it is inappropriate to give details of this work. Synge[20] has also used it for the fractionation of partial hydrolysis mixture of gramacidin.

A review with color photographs of column, and paper strip and sheet chromatograms has recently been published[29] and the literature to 1944 has been reviewed.[30]

IDENTIFICATION OF SIMPLE PEPTIDES

Consden, Gordon, and Martin[21] found that simple peptides will, in general, run on the chromatogram as satisfactorily as amino acids. However, both because of their large number and also because the R_F values are much more sensitive to slight changes in solvents, etc., it was not practicable to identify them by their R_F values alone. The authors therefore developed the technique of washing them off the chromatogram and hydrolyzing, and deaminating and hydrolyzing them, and then identifying the products on other chromatograms.

If the mixture was initially somewhat complex, duplicate two-dimensional chromatograms were run and one of these was treated with ninhydrin in the usual way. This reveals the full picture. The other was treated with about one-tenth the usual amount of ninhydrin, a quantity insufficient to destroy more than a small fraction of the peptide but enough to give an indication of the position of the stronger spots. The positions of the weaker spots shown on the normally treated sheet could usually be interpolated from the stronger spots. The pieces were cut out in this form , and the flat end was applied to wet filter paper overhanging a water-filled trough. To the dependent point of the piece cut from the chromatogram, a capillary tube was adjusted. In the course of about an hour, 50-100 microliters ran into the capillary and this was found to be sufficient to wash about 10 cm.^{2} of paper.

The liquid in the capillary was then transferred to a strip of polythene on which it formed a drop with no tendency to spread. The water was removed in a vacuum desiccator and a drop of 6N HCl added, the HCl solution was taken back into the capillary, the ends sealed, and put in an oven at 100°C overnight for hydrolysis. After removal of the HCl in the desiccator, again on a polythene strip, the amino acids were transferred to a paper chromatogram with water. If the peptide was to be deaminated, this was done by exposing it in 6N HCl on polythene to nitrous fumes.

Using this technique, Consden, Gordon, Martin, and Synge[22] studied the partial HCl hydrolysis products of gramicidin *S*. This has been shown by Synge[23] to consist of equimolecular proportions of the five amino acids, valine, ornithine, leucine, phenylalanine, and proline, and to be a cyclopeptide since there was only one free amino group and no free carboxyl group. The free amino group was shown by Sanger[24] to be the side chain amino of the ornithine. Among the partial hydrolysis products were identified valylornithine, ornithyl leucine, leucyl phenylalanine, phenylalanyl proline, and two tripeptides (not rigorously identified) phenylalanyl prolylvaline, and valyornithyl leucine.

Control chromatograms made up with the five amino acids and synthetic valylornithine, ornithyl leucine, leucyl phenylalanine, phenylalanyl proline, and prolyl valine gave a picture expected from that of the partial hydrolysis mixture, with the addition of the prolyl valine spot, and the synthetic mixture added to the partial hydrolysate showed only the expected number of spots. Prolyl valine proved to be more readily hydrolyzable than the other peptides, but this may not be the full explanation of the failure to observe it. These findings together with x-ray data from Dr. Crowfoot, consistent with Gramicidin-*S* being either a cyclopentapeptide or a cyclodecapeptide with a center of symmetry, show that the molecule consists of the amino acids, valine, ornithine, leucine, phenylalanine, and proline, united in that sequence once, or twice repeated, in a ring.

Consden, Gordon, and Martin[25,26] have examined hydrolysates of wool in a similar manner. The mixture was so complicated that little could be done unless separation by other means preceded chromatography. They therefore used ionophoresis as a first step and examined the fractions thus obtained chromatographically. Confirmation of their findings has to await synthesis of the many peptides encountered. Of particular interest, however, is the finding of several peptides of two united polar amino acids, and of two united non-polar amino acids, in contradiction to Astbury's hypothesis of the structure of keratin.

QUANTITATIVE ANALYSIS OF AMINO ACIDS BY PAPER CHROMATOGRAPHY

Martin and Mittelmann[27] attempted, by several methods, to render the paper chromatograms quantitative. The great stumbling block is, of course, the small quantities that can be handled, and presence on the paper of not negligible quantities of soluble materials other than the amino acids.

The method they found most promising was a modification of Pope and Stevens's method of forming the copper complex of the amino acid and then estimating the copper in this. So far, only simple mixtures of amino acids have been analyzed which could be separated in a one-dimensional chromatogram. The technique was as follows: On a wide sheet of paper, well-separated spots of a control mixture and the mixture to be analyzed were alternately placed and the chromatogram developed. The chromatograms at either edge of the sheet were cut off and developed with ninhydrin, and then used as guides for cutting up the rest of the sheet. The pieces containing the separated amino acids were then, after cutting into small pieces, stood in a suspension of copper phosphate in a borax phosphate mixture, in centrifuge tubes. After standing overnight in an icebox, the copper phosphate and paper were centrifuged down and the copper in the supernatant liquid estimated by the polargraph, after the addition of sodium sulphite. The amounts of the control mixture put on the paper were graded so that the best straight line could be drawn, on a graph showing current/concentration for each amino acid. From these graphs, the amounts present in the unknown could be inferred.

So far, the only natural product analyzed has been germicidin-*S*. Synge has shown this to contain equimolecular proportions of the five amino acids. The amounts found by Martin and Mittelmann are shown in TABLE 6. Though the accuracy of these results is less than that obtainable by the silica method using acetyl amino acids, and less also than that obtained by the best of the microbiological methods, less than one milligram of material was consumed in the whole analysis. Work on this method is at present interrupted, and it is too early, as yet, to decide upon its final usefulness.

TABLE 6

Amino acid	Solvent	Amount in each spot 10^{-7} mols.
Valine	*S*-Collidine	1.65
Proline		1.62
Ornithine		1.55
Leucine	Benzylalcohol	1.53
Phenylalanine		
Amount expected on Synge's formula: 1.64 × 10^{-7} mols.		1.67

BIBLIOGRAPHY

1. **Synge, R. L. M.**
1939. Biochem. J. **33**: 1913, 1918, 1924, 1931.
2. **Martin, A. J. P., & R. L. M. Synge**
1941. Biochem. J. **35**: 91.
3. **Martin, A. J. P., & R. L. M. Synge**
1941. Biochem. J. **35**: 1358.
4. **Gordon, A. H., A. J. P. Martin, & R. L. M. Synge**
1944. Biochem. J. **38**: 65.

5. **Liddell, H. F., & H. N. Rydon**
1944. Biochem. J. **38**: 68.
6. **Tristram, G. R.**
1946. Biochem. J. **40**: 721.
7. **Isherwood, F. A.**
1946. Biochem. J. **40**: 688.
8. **Hanby, W. E., & H. N. Rydon**
1946. Biochem. J. **40**: 297.
9. **Elsden, S. R.**
1946. Biochem. J. **40**: 252.
10. **Ramsey, L. L., & W. I. Patterson**
1945 J. Assoc. Off. Agric. Chem **28**: 641.
1946. J. Assoc. Off. Agric. Chem. **29**: 337
11. **Sanger, F.**
1945. Biochem. J. **39**: 507.
12. **Fischbach, H., M. Mundell, & T. E. Eble**
1946. Science **104**: 84.
13. **Gordon, A. H., A. J. P. Martin, & R. L. M. Synge**
1943. Biochem. J. **37**: Proc. xiii.
14. **Rheinboldt, H.**
1925. In: **Houben, J.** Die Methoden der Organischen Chemie (3rd ed.) **1**: 291. Georg Thieme. Leipzig.
15. **Consden, R., A. H. Gordon, & A. J. P. Martin**
1944. Biochem. J. **38**: 224.
16. **Consden, R., A. H. Gordon, A. J. P. Martin, O. Rosenheim, & R. L. M. Synge**
1945. Biochem. J. **39**: 251.
17. **Dent, C. E.**
1947. Biochem. J. **40**: Proc. xliv; also **41**: 240.
1946. Lancet **251**: 637.
18. **Partridge, S. M.**
1946. Nature **158**: 270.
19. **Chargaff, E.**
Private communication. (See **Fischer, E., & E. Chargaff.** 1947. J. Biol. Chem. **168**: 781.)
20. **Fink, R. M., & K. Fink**
Private communication.
21. **Consden, R., A. H. Gordon, & A. J. P. Martin**
1947. Biochem. J. **41**: 590.
22. **Consden, R., A. H. Gordon, A. J. P. Martin, & R. L. M. Synge**
1947. Biochem. J. **41**: 596.
23. **Synge, R. L. M.**
1945. Biochem. J. **39**: 363.
24. **Sanger, F.**
1946. Biochem. J. **40**: 261.
25. **Consden, R., A. H. Gordon, & A. J. P. Martin**
To be published.
26. **Martin, A. J. P.**
1946. Symposium on Fibrous Proteins. Soc. of Dyers & Colourists.
27. **Martin, A. J. P., & Rose Mittelmann**
To be published.
28. **Edman, P.**
1946. Arkiv. Kemi, Mineral. & Geol. (A 22) **3**.
29. **Martin, A. J. P.**
Endeavour **6**: 21. 1947.
30. **Martin, A. J. P., & R. L. M. Synge**
Advances in Protein Chemistry. **II.** Academic Press, Inc. 1946.

PARTITION CHROMATOGRAPHY OF AMINO ACIDS ON STARCH

By Stanford Moore and William H. Stein

The Rockefeller Institute for Medical Research, New York, N. Y.

In the preceding paper of this conference, Dr. A. J. P. Martin[1] has described the investigations carried out in England which have led to the development of the technique known as partition chromatography. Our own experiments in this field resulted from an interest in the isolation and identification of amino acids and peptides in biological material. The recently developed principles of partition chromatography appeared to offer marked potentialities for this type of problem.

In considering the form of partition chromatogram most suitable for isolation work with amino acids and peptides, we were impressed by the demonstration, by Elsden and Synge,[2, 3] that raw potato starch was a satisfactory supporting medium for the aqueous phase. As the authors showed, starch columns possess the advantage that they can be applied directly to the fractionation of mixtures of free amino acids and peptides. The preliminary acetylation procedure required in the case of the silica gel columns, with its attendant drawbacks, is thus avoided. The same advantages indicated for starch hold also for the paper strip chromatogram.[4] Paper strips, however, handle only microgram quantities of material, whereas columns handle milligram quantities. From a column, therefore, sufficient material may be obtained to permit further investigation by conventional microchemical techniques.

In 1944, Synge[3] employed starch columns in experiments with hydrolysates of gramicidin. *n*-Butanol-water was used as the solvent system. The use of free amino acids and peptides, a major proportion of which are neutral substances, meant that visualization of the separation process as it occurred on the column could not be effected by means of a pH indicator, as it could be in the case of the fractionation of the acetylated amino acids on silica gel columns. The positions of amino acids after development of the starch column with butanol could be ascertained, however, by passing through a solution of ninhydrin in ether. From the positions of the resulting blue bands, the rates of movement of the amino acids on the column could be determined. For the fractionation of the unknown constituents of a partial hydrolysate of gramicidin, Synge collected the effluent in a series of fractions and tested each one qualitatively with ninhydrin paper. The volume of each fraction was adjusted in order to collect substances within a given range of distribution coefficients. On working up the fractions individually, several of the components of the partial hydrolysate were isolated and identified.

In adapting the starch column to our problems, we have made two additions to this procedure. Instead of collecting the effluent from the chromatogram in large fractions, the effluent has been divided into a regular series of very small fractions of known volume. These fractions have been analyzed quantitatively rather than qualitatively. The data thus obtained permit the construction of effluent concentration curves which reveal the detailed behavior and the full resolving power of the column.

A typical effluent concentration curve, showing the separation of phenylalanine, leucine, and isoleucine is illustrated in FIGURE 1. The advantages to be gained from employing precise curves of this type to visualize the results, have been well demonstrated in the counter-current distribution studies of Dr. Lyman C. Craig.[5, 6, 7] It is a pleasure to acknowledge many helpful and encouraging discussions with Dr Craig on the subject of these experiments.

In the experiment referred to in FIGURE 1, the solvent was 1:1 *n*-butanol:benzyl alcohol nearly saturated with water. The effluent was collected in 1-cc. fractions. The amino acid concentration is expressed in millimoles per liter. The manual collection of 1-cc. fractions is obviously too tedious an occupation. The performance of a large number of experiments of this type became a practical laboratory procedure only after the construction of a fully automatic fraction collector,[8] which is illustrated in FIGURE 2. The column, with a solvent reservoir, is mounted over a circular rack containing eighty test tubes. The sample of amino acids, or amino acid hydrochlorides, dissolved in one or two cc. of the organic phase, has been added previously to the top of the column, drained into the starch, and the reservoir filled with solvent. Columns of the size shown

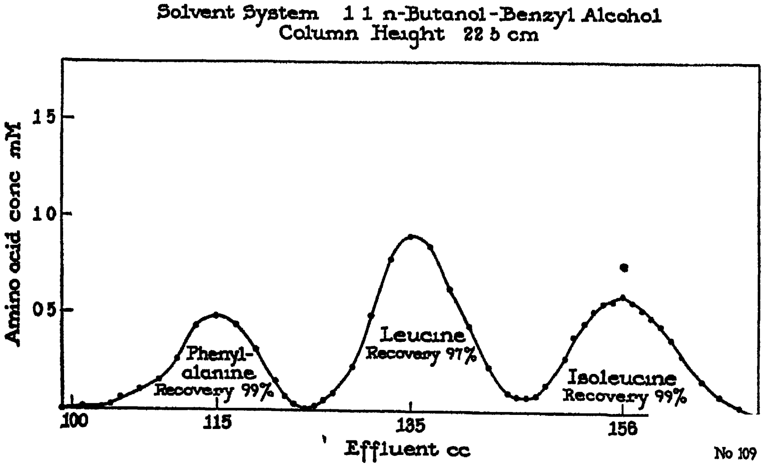


FIGURE 1. Separation of phenylalanine, leucine, and isoleucine.

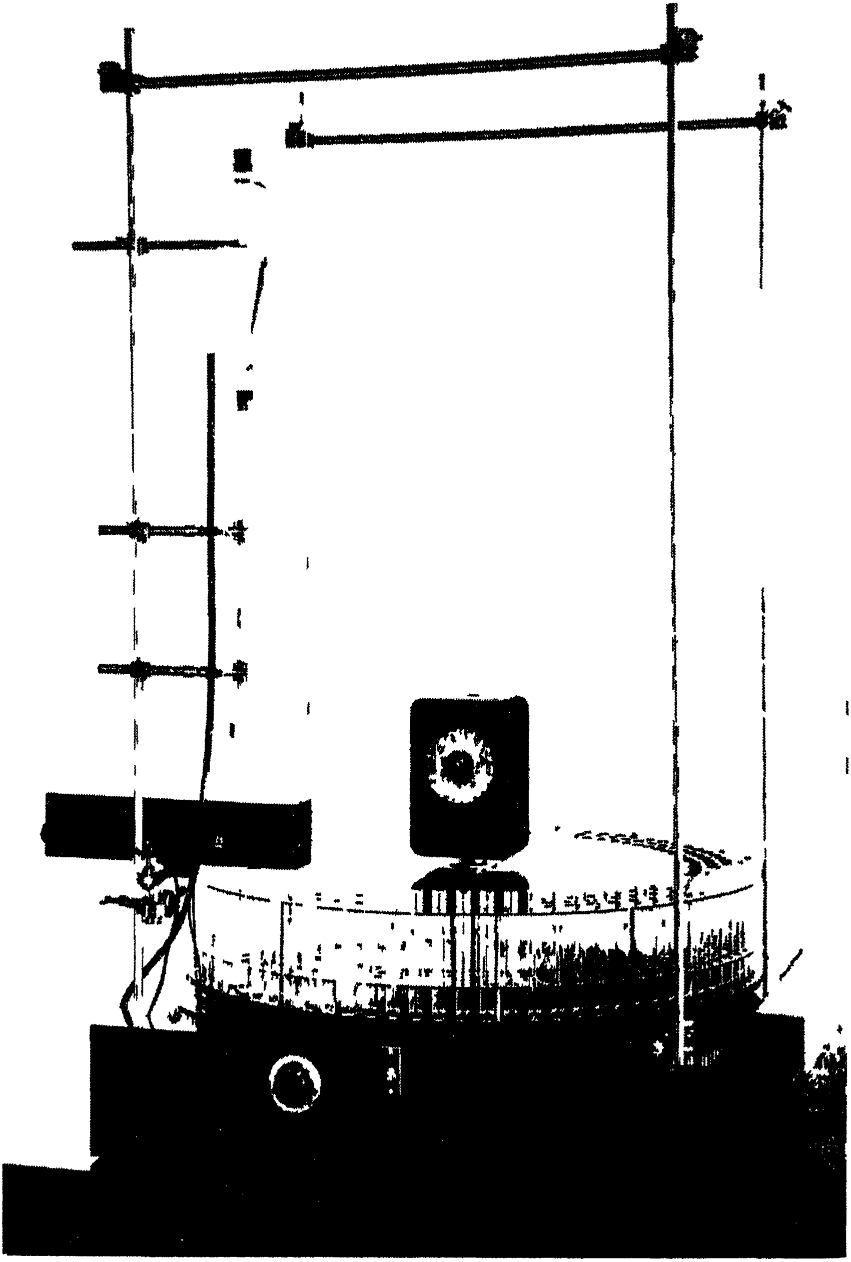


FIGURE 2. Automatic fraction collector for chromatographic analysis.

in FIGURE 2, about 2 cm. in diameter, are operated at a flow rate of about 5 cc. per hour, maintained either by gravity or by slight air pressure. As each drop emerges from the column, it intercepts a light beam focused on a photocell. The drops registered by the photocell are counted by an automatic-reset impulse counter. After a preset number of drops, corresponding to any desired volume from one drop to about 20 cc., the impulse counter resets itself to zero and closes the switch on a motor. The motor moves the rack 1/80 of a turn, bringing a new receiver under the column. The entire operation then starts over. A small funnel with a capillary tip blocks the mouth of each receiver, thus preventing evaporation. In practice, an individual experiment may run from 48 to 72 hours. The fraction collector permits day and night operation.

The next step involves the analysis of the numerous fractions collected. For this purpose, a quantitative colorimetric ninhydrin method has been developed.[9] Prior to quantitative analysis, however, the fractions are spot-tested on paper impregnated with ninhydrin. On the basis of the spot tests, the appropriate fractions can be chosen for quantitative analysis. For the determinations, small aliquots of the effluent fractions are pipetted into colorimeter tubes. A pipetting machine delivers, in the proper amount, the reagent which is stored under nitrogen in a reservoir attached to the machine. The colorimeter tubes are heated for 20 minutes in a boiling water bath, diluted to volume with a second pipetting machine, and read on a spectrophotometer. With this procedure, five hundred to one thousand analyses per week can be run conveniently by one individual.

The modified ninhydrin method gives reproducible results with most of the amino acids and with many peptides possessing a free α-amino group. The principal disadvantage of the method is that the amino acids do not all give the same intensity of color per mole. To the extent that the chromatogram accomplishes separation of the individual amino acids from one another, the colorimetric method is capable of yielding quantitative values for any given amino acid.

Another disadvantage of the ninhydrin method is that it determines only compounds possessing amino groups. Continuous recording refractometry would have more general applicability. The sensitivity of the colorimetric ninhydrin method, however, which is 1 ppm. of amino acid in aqueous or alcoholic solution, is about 150 times that of refractometric measurements with an accuracy of $2 \times 10^{-5}\ n_D$. The combination of simplicity and sensitivity in the analytical method has been of importance in the experiments with partition chromatograms. Although the starch columns can handle milligram quantities, initially in fairly concentrated solution, the concentrations of components in the effluent progressively decrease as the fractionation process is extended to larger effluent volumes.

In returning to a consideration of FIGURE 1, it will be noted that, for cases where the procedure has been well worked out, the integration of the curves gives quantitative recovery for the components. Before a curve as satisfactory as this one could be obtained, many variables had to be brought under control. An early experiment with the same three amino acids and the same solvent system is shown in FIGURE 3. The irregularity and poor resolution in this case may be ascribed to several causes, among which are poor packing of the column, too rapid a flow rate, and suspended water droplets in the organic solvent.

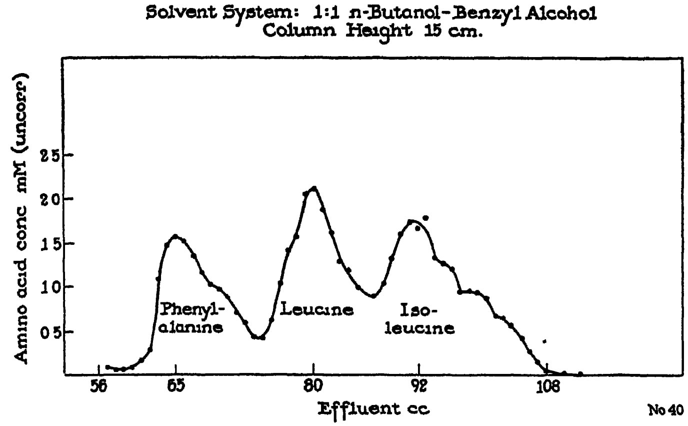


FIGURE 3. Results on an unsatisfactory column.

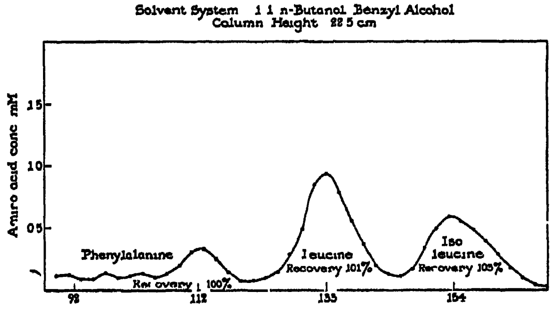


FIGURE 4. Anomalous behavior of the first band.

A difficulty which bothered us for some time is illustrated in FIGURE 4. The *first* band, phenylalanine in this case, exhibits an anomalous behavior. There is a long segment of the band running 20 cc. ahead of the peak. The second and third peaks are symmetrical. If an aliphatic amine or 8-hydroxyquinoline is allowed to run through the column ahead of the first amino acid band, the effect can be eliminated. Both of these substances are capable of forming complexes with metals. The use of

8-hydroxyquinoline has proved to be essential for the procurement of satisfactory results with all of the amino acid bands, when samples of starch are used containing significantly higher traces of inorganic impurities than the starch used in this experiment.

It is possible that the anomaly is similar to that noted by Consden, Gordon, and Martin with the paper-strip chromatogram, and ascribed to the presence of traces of copper.[4] Cupron, which abolished the effect with paper, is, however, ineffective on the starch column.

The behavior of the faster-moving amino acids in three solvent systems is given in FIGURE 5. The solvents are butanol-water, butanol-0.5N HCl, and 1:1 butanol:benzyl alcohol-water. In the first curve, which represents an early experiment with butanol-water, the shoulder on the leucine peak is the result of the effect of metals mentioned above. In the second curve, illustrating an experiment in which butanol-0.5N HCl is the solvent system, the sequence of the amino acid bands is altered. In this instance, the amino acid hydrochlorides are fractionated. In reference to the third

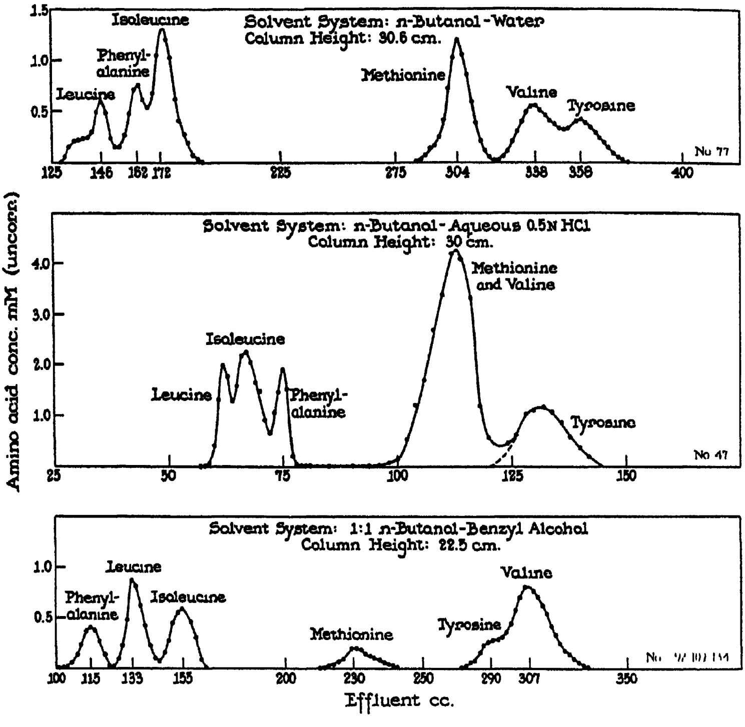


FIGURE 5. Separation of phenylalanine, leucine, isoleucine, methionine, valine, and tyrosine.

curve, as found by Consden, Gordon, and Martin on paper strips, the use of 1:1 butanol:benzyl alcohol serves to increase the band rates of the aromatic amino acids relative to the rates of the aliphatic ones. Thus, phenylalanine moves ahead of leucine and isoleucine, and tyrosine emerges ahead of valine. A more complete resolution of these three amino acids is attained in the butanol:benzyl alcohol system than in other solvent mixtures which have been studied. The recoveries of methionine, however, run low in butanol:benzyl alcohol experiments. Oxidation of the methionine sulfur by traces of impurities in the benzyl alcohol is believed to be responsible for the low yields. Satisfactory recoveries are obtained with methionine in plain butanol. It should be mentioned that tryptophane, if present, would emerge near the phenylalanine band in all three solvents, and would show up in the valley between phenylalanine and leucine in butanol:benzyl alcohol. These amino acids can be separated from one another by re-chromatographing the mixture on a separate column, as will be mentioned later. In acid hydrolysates of proteins, the problem does not arise, since tryptophane is decomposed during the hydrolytic process.

Before discussing the amino acids which appear later in the amino acid "spectrum," some of the fundamental theoretical points raised by the results shown in FIGURE 5 merit consideration. In its original formulation, "partition chromatography" was conceived to be a technique which permitted counter-current liquid-liquid extraction to be carried out as a simple column filtration with one of the phases immobilized on an *inert* support. The data in TABLE 1 serve to indicate the extent to which this conception of the operation of the column agrees with the experimental facts. In this table, the distribution coefficients of several amino acids are compared with their rates of movement on starch columns. In the first section of the table, the distribution coefficients of the amino acids in butanol-water as determined in the separatory funnel are listed in decreasing order.[8] The distribution coefficients from the rates of movement of the amino acids on the column have been calculated in each case by the method of Martin and Synge.[10] The equations employed in these calculations were derived on the assumption that the amino acids are being separated by a liquid-liquid distribution process. For leucine, isoleucine, and valine the distribution coefficients calculated from the column are somewhat higher than those found in the separatory funnel, but the relative distribution coefficients are similar in the two cases. These data support those reported by Synge,[3] who found that the distribution coefficients of leucine, valine, and alanine, calculated from starch columns, agreed fairly well with the values determined by England and Cohn[11] by the separatory funnel technique. For the simple aliphatic monoamino acids, therefore, with butanol as the solvent, the agreement between experiment and theory would make it appear that starch is serving as a relatively inert support for the aqueous phase.

TABLE 1

COMPARISON OF DISTRIBUTION COEFFICIENTS OF SEVERAL AMINO ACIDS WITH THEIR RATES OF MOVEMENT ON STARCH COLUMNS

Solvent System: *n*-Butanol-Water

Amino Acid	Tryptophane	Phenylalanine	Leucine	Isoleucine	Tyrosine	Methionine	Valine
Distribution coefficient in separatory funnel *	0.56	0.240	0.181	0.132	0.117	0.091	0.056
Distribution coefficient calculated from the column †	ca. 0.22	0.20	0.22	0.18	0.080	0.096	0.086
Order of emergence from the column	2	3	1	4	7	5	6
Order of band rates on paper ‡	4	1	2	3	5	6	7

Solvent System: 1:1 *n*-Butanol: Benzyl Alcohol-Water

Amino Acid	Tryptophane	Phenylalanine	Leucine	Tyrosine	Isoleucine	Methionine	Valine
Distribution coefficient in separatory funnel *	0.60	0.270	0.123	0.104	0.097	0.082	0.044
Distribution coefficient calculated from the column †	0.20	0.21	0.18	0.072	0.15	0.094	0.068
Order of emergence from the column	2	1	3	6	4	5	7
Order of band rates on paper ‡	3	1	2	6	4	5	7

* Distribution coefficient $= \frac{\text{conc. in non-aqueous phase}}{\text{conc. in aqueous phase}}$, determined with 0.01 M amino acid solutions at 25°C. Accuracy ca. ± 2%[8].

† Calculated by the method of Martin & Synge[10].

‡ From data of Consden, Gordon, & Martin[4].

However, when the distribution coefficients in butanol-water of all of the amino acids given in TABLE 1 are determined, a number of deviations from the theory become apparent. If the starch column were operating as a true liquid-liquid distribution system, the amino acids should, in each case, emerge in the order of the distribution coefficients as listed in decreasing order. The actual orders of emergence are those shown in FIGURE 5 and listed in the table. The scrambling of the order results from the behavior of phenylalanine, tryptophane, and tyrosine, which do not travel on the column at the rate which would be predicted from the distribution coefficients. Tryptophane should be as far ahead of phenylalanine as isoleucine is ahead of valine; actually, it is behind leucine. Phenylalanine should be well ahead of leucine, instead of behind it, and tyrosine should be between isoleucine and methionine, whereas, actually, it emerges last. Furthermore, the effect upon the aromatic amino acids is not peculiar to starch. As may be seen from the data of Consden, Gordon, and Martin, listed in TABLE 1, similar anomalies are noted with cellulose in the paper-strip chromatogram. Tryptophane, for example, is markedly out of order in this instance.

The data for 1:1 butanol:benzyl alcohol are given in the second section of TABLE 1. In the case of the paper-strip chromatogram, the deviations from the expected order are more marked than is the case in butanol. The same type of deviation from theory has also been demonstrated for the butanol-0.5N HCl system, for which the orders of emergence are given in FIGURE 5.

The conclusion to be derived from these data is unequivocal, namely, that starch does not serve merely as an inert support for the aqueous phase.

In this connection, the following simple experiments have been carried out in which no organic solvent whatever has been employed. An aqueous solution of glycine and alanine has been passed through a column composed of only starch and water (FIGURE 6). The separation is quantitative. The experiment has been repeated with leucine, phenylalanine, and tryptophane (FIGURE 7). Again, with no organic solvent present, complete separation is obtained. That this effect is not pH-sensitive, is demonstrated by the dotted curve in the figure. This second curve was obtained by fractionating the hydrochlorides of the amino acids on a column made up with starch and aqueous 0.1N hydrochloric acid.

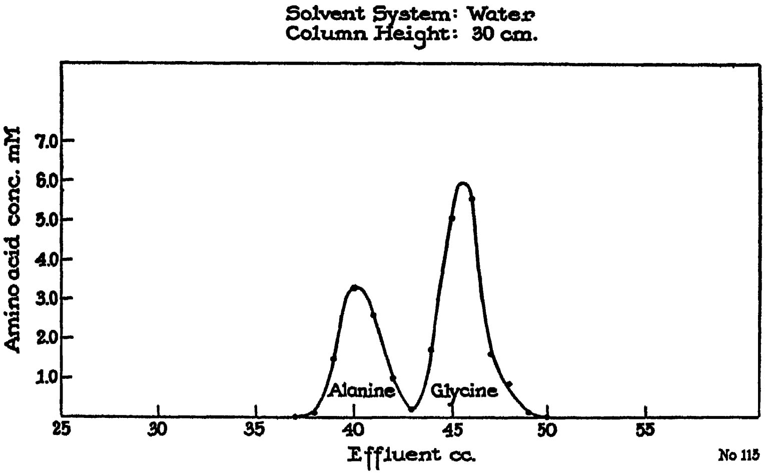


FIGURE 6. Separation of glycine and alanine on a water-starch column.

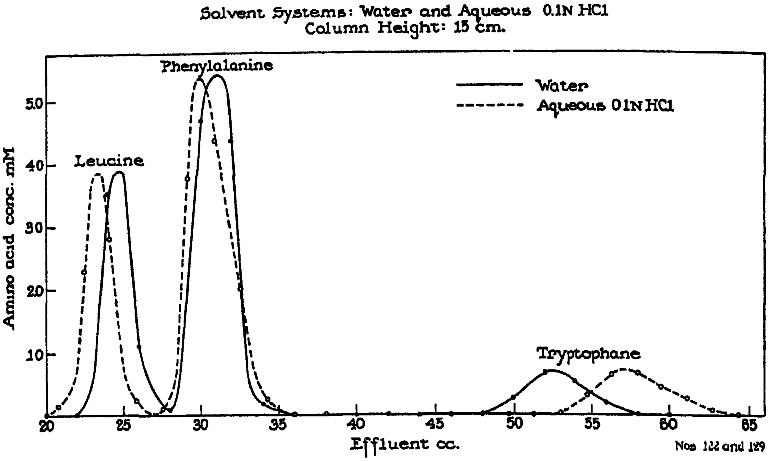


FIGURE 7. Separation of leucine, phenylalanine, and tryptophane.

In this instance, the fractionation can be a useful one for separating tryptophane from phenylalanine and leucine, which, as was mentioned earlier, is a separation difficult to accomplish with organic solvents. Thus, some of the experiments which were run primarily for their theoretical interest led to the realization that useful results with amino acids may be obtained on starch columns using water-miscible solvents. This point will come up again in the discussion of the chromatography of the basic amino acids. There are indications from the work of Consden, Gordon, and Martin that amino acids can also be fractionated on paper strips with the aid of solvents miscible with water.[4]

From the foregoing, it would appear that the starch column possesses properties characteristic of both an adsorption chromatogram and a partition chromatogram. In view of the known adsorptive properties of the starch granule,[12] this is not an unexpected finding. The data indicate clearly that further work is necessary in order to elucidate the role of adsorption in the performance of both the starch and the cellulose chromatograms.

Two further points which bear upon the theoretical interpretation warrant mention. The first point concerns the symmetry of the effluent concentration curves. The symmetry is indicative of a fairly linear isotherm, irrespective of the definition of the process by which the fractionation is attained. The second point of importance is the reproducibility of the absolute position in the "spectrum" of a given amino acid band. When butanol-benzyl alcohol is employed as the solvent, the amino acid valine, for example, in the presence of twenty other components, emerges at the same position as it does when it is the only amino acid on the column. The same holds true for all the other amino acids which have been tested thus far. This situation is in contrast to that frequently encountered in adsorption chromatography, where the behavior of a given substance is markedly altered by the presence of other components.[13, 14]

Before leaving the discussion of the theory of chromatography on starch, it is of interest to consider the order of magnitude of the resolving power of the starch column. The case of leucine and isoleucine in 1:1 butanol:benzyl alcohol is an instance where the ratios of the two distribution coefficients calculated from the column and measured in the separatory funnel are similar. For purposes of comparison, without making any assumptions as to the mechanism of action of the column, it is possible to calculate how many "plates" would be required to achieve, by liquid-liquid distribution in a series of separatory funnels, the same degree of separation of these two amino acids attained on the column. The methods of calculation of Martin and Synge[10] and of Craig,[5] taken in conjunction with our data, point to the fact that the starch column is performing a task which would require several thousand separatory funnel "plates."

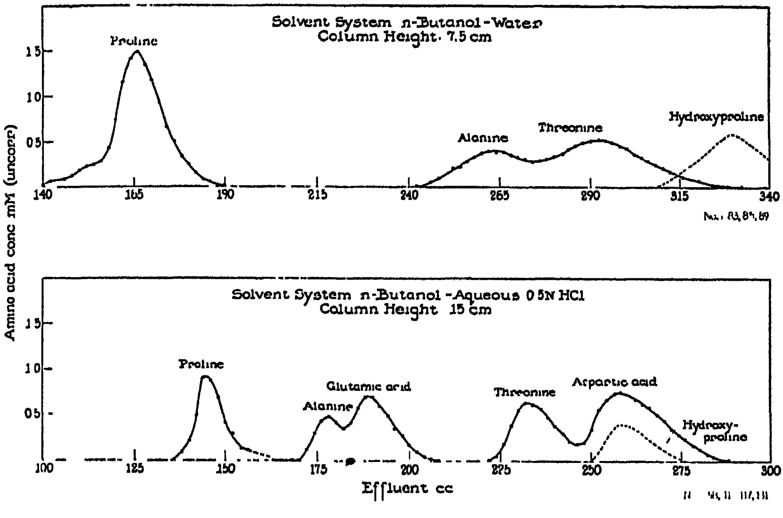


FIGURE 8. Separation of amino acids with band rates in the range of that of alanine.

The comparison of the performance of the column with that of a several thousand "plate" system is a realistic one for revealing, qualitatively, the extremely high resolving power inherent in the starch chromatogram and for indicating the position of this technique among the tools available to the organic chemist for the study of complex multi-component systems.

In returning to the discussion of the amino acid "spectrum," it should be stated that there has been no opportunity, as yet, to study the slower-moving bands as thoroughly as the phenylalanine-through-valine group has been examined. The results which follow are of a more preliminary nature and are presented to show that it has proved possible to cover the whole spectrum, although the degree of resolution is not satisfactory in all instances. Two curves illustrating the behavior of a number of the amino acids falling in the next group to emerge from the column are given in FIGURE 8. A shorter column is used, and the incompletely separated faster bands, if present, are accumulated in a fore-fraction. In butanol-water, proline is the next amino acid to emerge after tyrosine and valine, and it is well separated from alanine, threonine, and hydroxyproline. A different picture is obtained employing the acid solvent mixture, butanol-0.5N HCl. Alanine is pulled well apart from threonine. Hydroxyproline has about the same rate of travel as aspartic acid. Most of these overlaps are resolvable by re-chromatographing the zones in question. It is probable that further study of solvent systems will lead to results which will reduce the number of instances where this would be necessary.

Serine and glycine are the most slowly-moving members of the monoamino monocarboxylic acid series. To speed up their rate of movement, *n*-propanol can be used in preference to *n*-butanol (FIGURE 9). In neither of these solvents, however, has complete resolution been obtained.

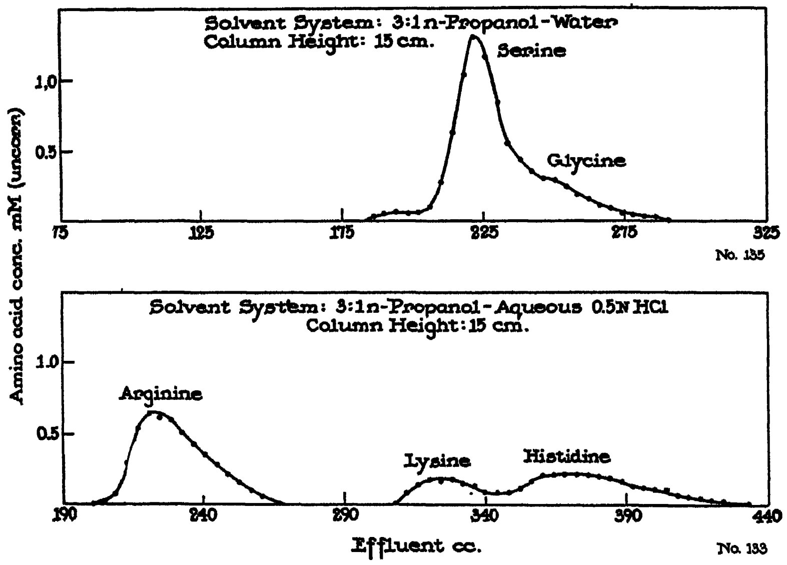


FIGURE 9. Separation of serine, glycine, and the basic amino acids.

The basic amino acids present a special problem. In butanol-water and in butanol-0.5N HCl, arginine, lysine, and histidine do not move at an appreciable rate. Even in plain water, arginine and lysine are bound strongly near the top of the column. The addition of HCl to the water causes all three basic amino acids to move rapidly. The combination of aqueous HCl and propanol provides a solvent mixture which gives a satisfactory rate of travel for the basic amino acids and yields reasonably complete separation (FIGURE 9). This preliminary experiment is a further example of the use of water-miscible instead of water-immiscible solvents in the starch chromatogram.

Relatively few experiments have been carried out, to date, on the behavior of peptides on the column. One example can be given of the separation of the isomeric dipeptides leucylglycine and glycylleucine (FIGURE 10).

The results of an experiment on a hydrolysate of β-lactoglobulin are given in FIGURE 11, as an example of one of the problems to which this technique can be applied. The first six amino acids in the spectrum appear in the usual order. In this mixture of more than twenty components, the positions of the bands are the same as when the amino acids

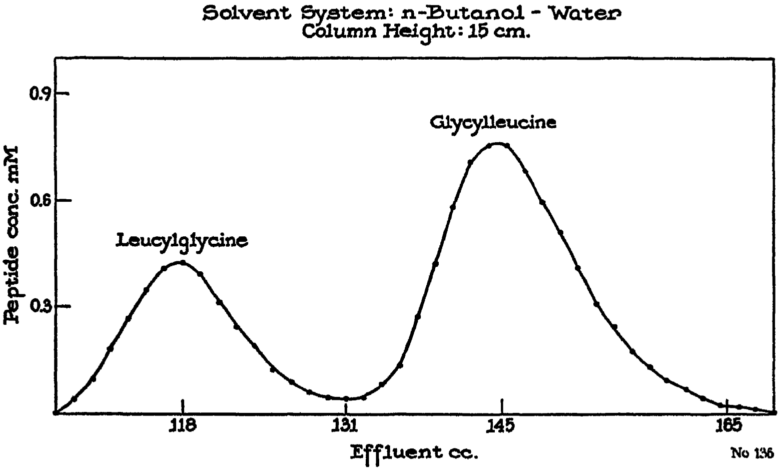


FIGURE 10. Separation of leucylglycine and glycylleucine.

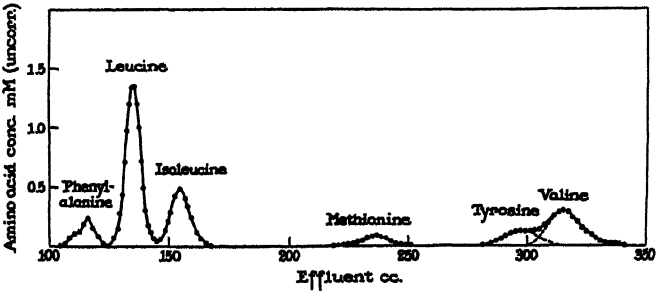


FIGURE 11. Separation of amino acids in a hydrolysate of β-lactoglobulin.

are run individually. Integration of the curves gives the following values for the amino acid composition of β-lactoglobulin (gm. per 100 gm. of protein)[8]: phenylalanine, 3.78; leucine, 15.5; isoleucine, 5.86; tyrosine, 3.64; and valine, 5.62. The values are in good agreement with the results reported by Brand *et al.*,[15] with the exception that the starch chromatogram indicates a lower value for the isoleucine content. The results which have been obtained with this method of fractionation support the conclusion that chromatography of free amino acids on starch columns, as introduced by the work of Martin and Synge and Elsden, and as extended in the experiments outlined above, is capable of yielding sound quantitative data as well as qualitative pictures of the composition of mixtures of amino acids and related compounds. The studies are still in their relatively early stages, and much is yet to be learned concerning both the potentialities and the limitations of the method in its application to biochemical problems.

In conclusion, the authors wish to acknowledge the valuable assistance of Miss Enid Mellquist and Mr. H. R. Richter in the course of this work, and the microanalytical aid of Dr. Adalbert Elek of the Institute.

BIBLIOGRAPHY

1. **Martin, A. J. P.**
   1947. Ann. N. Y. Acad Sci. **49**(2): 249-264.
2. **Elsden, S. R., & R. L. M. Synge**
   1944. Biochem. J. **38**: Proc. ix.
3. **Synge, R. L. M.**
   1944. Biochem. J. **38**: 285.
4. **Consden, R., A. H. Gordon, & A. J. P. Martin**
   1944. Biochem. J. **38**: 224.
5. **Craig, L. C.**
   1944. J. Biol. Chem. **155**: 519.
6. **Craig, L. C., C. Golumbic, H. Mighton, & E. Titus**
   1946. Science **103**: 2680.
7. **Hogeboom, G. H., & L. C. Craig**
   1946. J. Biol. Chem. **162**: 363.
8. **Stein, W. H., & S. Moore**
   1947. Unpublished data.
9. **Moore, S., & W. H. Stein**
   1947. Unpublished data.
10. **Martin, A. J. P., & R. L. M. Synge**
    1941. Biochem. J. **35**: 1358.
11. **England, A., & E. J. Cohn**
    1935. J. Am. Chem. Soc. **57**: 634.
12. **Strain, H. H.**
    1942. Chromatographic Adsorption Analysis. Interscience Publishers, Inc. New York.
13. **Zechmeister, L., & L. Cholnoky**
    1944. Principles and Practice of Chromatography. John Wiley & Sons, Inc. New York.
14. **Cannan, R. K.**
    1946. Ann. N. Y. Acad. Sci. **47**(2): 135.
15. **Brand, E., L. J. Saidel, W. H. Goldwater, B. Kassell, & F. J. Ryan**
    1945. J. Am. Chem. Soc. **67**: 1524.

A REVIEW OF FRACTIONATION OF MIXTURES BY FOAM FORMATION

By LEO SHEDLOVSKY
Colgate-Palmolive-Peet Company, Jersey City, N. J.

Chromatographic methods employing solid adsorbents are now frequently used, but the possibilities of fractionations and purifications with foams (at air-liquid interfaces) and emulsions (at liquid-liquid interfaces) have hardly been investigated.

It has been known for some time that the components of certain solutions are adsorbed by the foams which these solutions can form, as well as at liquid-liquid interfaces. Gibbs' Adsorption Equation* indicates increased concentration of a single surface active material at the interface by reversible adsorption. In a number of cases, the measured adsorption is equal to or many times the calculated amount.

MEASUREMENT OF ADSORPTION AT LIQUID-LIQUID AND GAS-LIQUID INTERFACES

W. C. M. Lewis[1] estimated the adsorption of sodium glycocholate, sodium oleate, congo red, and methyl orange for aqueous solutions against hydrocarbon oil and found it to be about sixty times the calculated value. The adsorption was about 5×10^{-6} gm. per $\mathrm{cm.}^2$

Donnan and Barker[2] found the adsorption at gas-liquid interfaces, as indicated by surface concentration, to be of the order of 10^{-7} gm. per $\mathrm{cm.}^2$ for dilute solutions (0.002 to 0.008 per cent) of nonylic acid. These values agree in order of magnitude with those calculated from Gibbs' Equation.

McBain and DuBois[3] measured adsorption at the surface of aqueous solutions for isoamyl alcohol, acetic, butyric, caproic and nonylic acids, phenol, *p*-toluidine, resorcinol, thymol, and camphor. They found that gas bubbles passing through a solution carry from two to eight times as much solute as is predicted by Gibbs' Equation.

$\Gamma = -\frac{a}{RT} \times \frac{d\gamma}{da}$, where Γ is excess solute per unit surface, a = activity of solute, and $\frac{d\gamma}{da}$ = rate of change of surface tension with activity. Where activity coefficients do not change with concentration, concentrations are used instead of activities.

[1] Lewis, W. C. M. Phil. Mag. 15: 499-526. 1908.

[2] Donnan, F. G., & J. T. Barker, Proc. Roy. Soc. London A 85: 557-73. 1911.

[3] McBain, J. W., & R. DuBois. J. Am. Chem. Soc. 51: 3534-3549. 1929.

PURIFICATION OF MIXTURES BY SELECTIVE ADSORPTION IN FOAM OR EMULSION

The above values show the extent of adsorption for a number of substances. For mixtures, fractionation of the constituents can be obtained by selective adsorption at air-liquid interfaces (foam) or liquid-liquid interfaces (emulsions). This suggests the use of foams for the removal of positively adsorbed substances from solution. However, very little quantitative work along these lines has been reported, and for adsorption in emulsions even less data are available.

In the following discussion, the recent papers[4-10] which indicate some of the factors involved in such separations will be considered.

Miles and Shedlovsky[4, 5, 6] have shown the effects of small amounts of certain impurities which lead to minima in surface tension-concentration curves. Such minima can be attributed to the presence of at least two surface active constituents in the same solution. Selective adsorption at air-liquid interfaces (foam) or benzene-liquid interfaces (emulsion) have indicated these minima to occur at bulk concentrations where the relative surface concentrations of the component present in smaller amount was at a maximum.

As an example of selective adsorption at air-liquid interfaces, the data given by Miles[5] may be considered. Known volumes of foam of uniform bubble size were produced by passing clean air, at constant pressure, through a single glass orifice into 50 ml. of solution. The interfacial area was approximately 2 x 10^4 cm.^{2} for a full cylinder of foam, as calculated from the bubble volume, which was estimated as 0.017 ml., and the cylinder volume of 1100 ml. After forming the foam, the solutions were brought to what was believed to be a steady state with respect to the newly created surface, by agitating the bulk liquid gently in contact with the interface for several minutes.

Curve 1 of FIGURE 1 represents the surface tension-concentration relationships for pure sodium lauryl sulfate which had been ether-extracted in the dry state for 36 hours in a Soxhlet apparatus. Curve 2 illustrates the minimum obtained when pure sodium lauryl sulfate is contaminated with 0.5 per cent dodecanol, on the total solids basis. Point A′ is the value obtained after forming a full cylinder of foam from 50 ml. of 0.015 molal solution (surface tension A) and then diluting to 0.0075 molal. Point B′ is the value obtained when the same procedure is followed starting with 0.010 molal, (surface tension B) and then diluting to 0.0075 molal after selective adsorption. Point C′ was obtained by the same steps as Points A′ and B′, except that the starting and final concentrations were

---

[4] Miles, G. D. & L. Shedlovsky. J. Phys. Chem. 48: 57. 1944.

[5] Miles, G. D. J. Phys. Chem. 49: 71. 1945.

[6] Shedlovsky, L. Ann. N. Y. Acad. Sci. 46(6): 429. 1946.

[7] Abribat, M. Actualités Scientifiques et Industrielles. Hermann et C^{ie}. Paris. 932: 128–175. 1942; C. R. 209: 244-6. 1939; U. S. Patent 2,313,007. Mar. 2, 1943.

[8] Marks & Clerk. Brit. Patent Appl. 27,886. Oct. 13, 1939.

[9] Schütz, F. Trans. Far. Soc. 42: 437–444. 1946. French Patent 835,439. Dec. 21, 1938.

[10] Bader, R. & F. Schütz. Trans. Far. Soc. 42: 571–76. 1946.

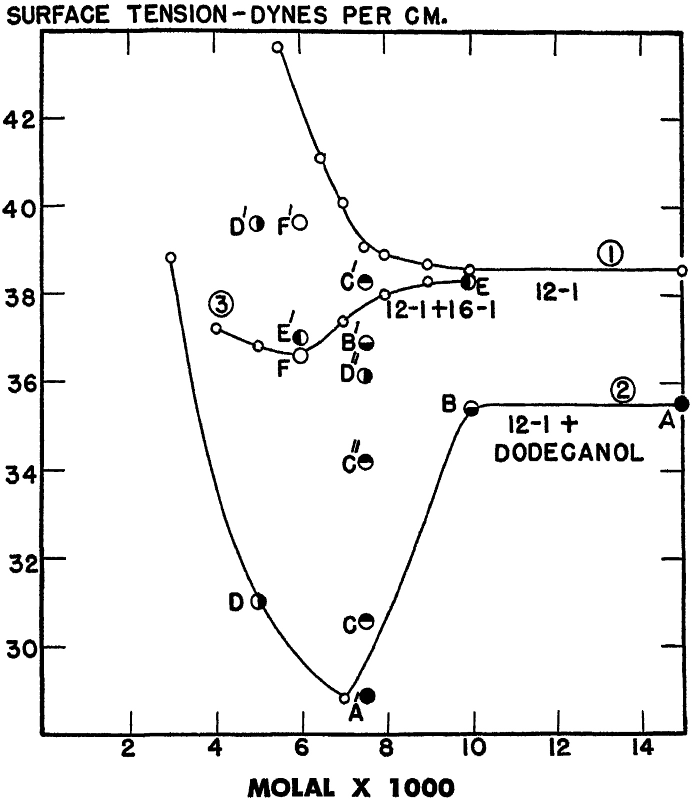


FIGURE 1. Effect of foam extraction on surface tension of sodium dodecyl sulfate (12-1) + 0.5% dodecanol and sodium dodecyl sulfate (12-1) + 1% sodium hexadecyl sulfate (16-1). (Miles [5]).

0.0075 molal. Point C″ was obtained in the same way as C′, except that one-third the volume of foam was employed for C″ as compared to C′. Fifty ml. of 0.005 molal solution, surface tension (D), was used to make a full cylinder of foam and the surface tension (D′) of the bulk liquid measured. On increasing the concentration to 0.0075 molal by the addition of pure sodium lauryl sulfate to this solution, a surface tension D″ was obtained. Point F′ was obtained by the same adsorption procedure, starting from a 0.006 molal solution with surface tension F.

The examples given are typical of other mixtures, and surface tension-concentration and interfacial tension-concentration curves are effective guides to the concentrations most suitable for purification by selective adsorption, as well as to show the presence of even minute amounts of certain surface active constituents. This gives criteria of purity not obtainable from the measurement of bulk properties. By studying both the surface tension-concentration and interfacial tension-concentration curves, using a non-aqueous liquid for the latter, it is sometimes possible to choose the method of purification to be used. If the surface tension-concentration curve shows a minimum, selective adsorption in foam is indicated, whereas, if the interfacial tension-concentration curve shows a minimum, selective adsorption in an emulsion should be tried. In either case, the most effective removal of the impurities is obtained at the concentration corresponding to the minimum. For example, we have found that addition of dodecanol to a solution of sodium dodecanol sulfate shows a pronounced minimum in the surface tension-concentration curve, but no minimum in the interfacial tension-concentration curve (solution against benzene). The dodecanol can be completely removed from this solution by the use of foam, whereas a benzene emulsion is not suitable for this purpose.

The procedure for separations with foams or emulsions will depend on whether it is desired to recover the material adsorbed in the foam or that remaining in the solution. For removal, from the solution, of materials which are selectively adsorbed in the foam, the method which we use at present consists of preparing a column of foam in a cylindrical pyrex jacketed tube 46 cm. long and 3.7 cm. in diameter, from another jacketed tube placed below the first one. The lower tube is 17 cm. long and 2.8 cm. in the widest portion. Filtered air is bubbled through a stainless steel perforated cup "spinnerette" (40 holes, 0.0032 inches in diameter). Surface tension measurements are made of the solution at various concentrations after it has been poured through the foam, and about 15 minutes are allowed for drainage. In some cases, repeating the operation of pouring the solution through the foam gives improved selective adsorption and a more complete approach to a steady state.

For selective adsorption at benzene-liquid interfaces, mentioned previously,[5] relatively small portions of the solution were gently agitated with the emulsion, whereupon samples of the bulk solution were withdrawn and their interfacial tensions compared with the values obtained for untreated solutions.

These selective adsorption experiments were made with mixtures where one constituent was present in small amounts. It is not known whether a similar procedure could be used satisfactorily for binary mixtures where both components were present in relatively large amounts. Furthermore, its applicability to multicomponent and more complex mixtures has not been investigated.

ABRIBAT'S METHODS FOR FRACTIONATION OF MIXTURES IN FOAM

M. Abribat[7] describes certain methods of fractionation of various materials using foam. He emphasizes that drying of foam by drainage is responsible for the enrichment of foams with respect to certain constituents in the solution, and he insists that the separation of numerous constituents is primarily a function of the difference in the flow characteristics at the gas-liquid interfaces and of the bulk liquid. Furthermore, he claims that this separation does not necessarily correspond to the surface activity of the constituents involved.

Procedure. The foam is made by bubbling air or nitrogen through a porous cup (E) at the bottom of a vertical glass tube (FIGURE 2). The gas bubbles should be as small as possible, so that they do not rise too rapidly, allowing adequate time for adsorption.

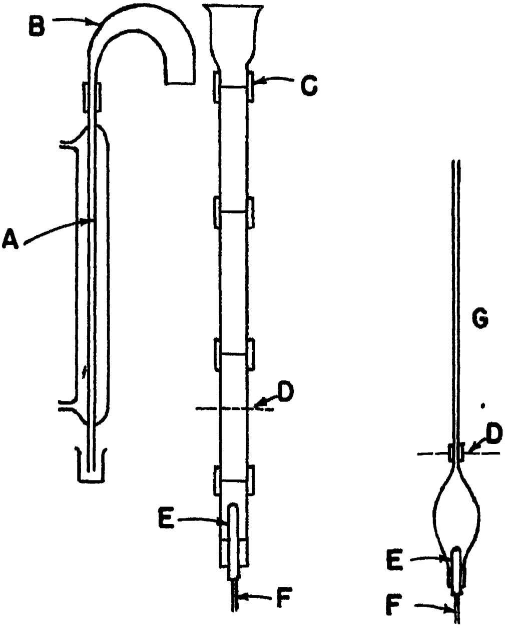


FIGURE 2. Foam fractionation apparatus, after Abribat[7]. A, defoaming tube; B, connecting tube; C, rubber sleeve; D, liquid level; E, porous cup; F, gas inlet; G, thin wall apparatus for solutions which form small amount of foam.

Since certain constituents of the original solution are concentrated in the foam, the amount of foam which can be formed from the solution decreases as it progressively loses those constituents which cause it to foam. In order to obtain effective separations, it is important to allow the foam to drain as much as possible. For some foams whose films have a high surface viscosity, such as protein foams, considerable liquid remains even after drainage has been essentially complete.

The height of the column of foam, H, from the original liquid level and the drop in the liquid level, h, are recorded as a function of time, with the gas bubbling at a fixed pressure (FIGURE 3). In general, H passes through a maximum for a constant gas pressure. These measurements are used by Abribat as a guide to the procedure to be followed in foam fractionation. He considers the type of data shown in FIGURE 3 as indicating

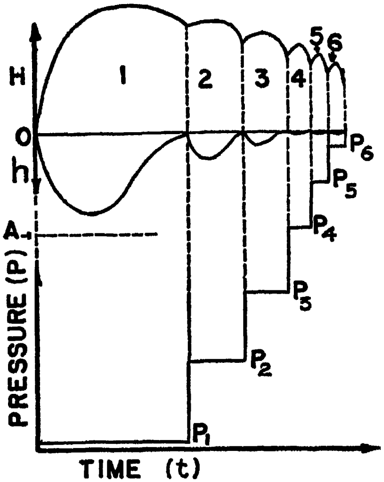


FIGURE 3. Schematic diagram for operation of foam fractionation, after Abribat[7]. *H*, height of column of foam; *h*, drop in liquid level.

relative drainage rates, but it seems that both the foam volume and the maximum total amount of liquid held in the foam, which is proportional to h, decrease for each increase in pressure at which the gas is bubbled.

To carry out an experiment, the bubbling of gas is first started at low pressures which are progressively increased until H is constant. The pressure should always be below the limit which corresponds to the appearance of eddy currents of bubbles in the foam. The porous cup should be at least 5 or 6 cm. below the liquid level for a tube of 5 cm. inside diameter.

After a few preliminary trials, a suitable pressure, P_1 is chosen (FIGURE 3). Then, t_1 is the time required to reach a steady state when foam breakdown is compensated by foam formation and H is constant. The pressure is increased to P_2 and the measurements of H and h against t are repeated.

The values of t become progressively shorter. After depletion of the solution of surface active constituents, and without diminishing the gas pressure, the sections of foam are removed. This can be done in a single operation if the apparatus is large enough. Otherwise, the sections of foam may be removed in steps corresponding to t_1, t_2, t_3, etc. For this purpose, the connecting tube (FIGURE 2, *B*) is placed over any one of the sections and the required foam fraction is collected in the defoaming tube.

Abribat's apparatus (FIGURE 2) is made with long cylindrical sections connected with ground joints or rubber sleeves, and the dimensions range from narrow columns several feet high to wide ones reaching a total of perhaps 20 feet. The size suitable for all operations cannot be specified, because this varies with the amount of foam formed. A relatively narrow tube is recommended for solutions which form only a small amount of foam, such as dilute solutions of proteins, virus, and serums. If a vertical tube with a sufficiently thin wall is used (FIGURE 2, *G*), the column may be cut into appropriate sections with a glass cutter at the end of the experiment. If necessary, the fractions obtained can again be treated separately by forming new foams, either after dilution or directly in a smaller apparatus. Other details of the apparatus are indicated in FIGURE 2. A jacketed apparatus is used to carry out experiments at various temperatures.

Effect of Concentration of Solutions and of Bulk and Surface Viscosity. According to Abribat, in many cases and especially for protein solutions, the fractionation by foam drainage is more complete for dilute solutions, as long as they still form rigid surface films. In this case, the difference between surface and bulk viscosity is greater than for more concentrated solutions. The addition of materials with high bulk viscosity is usually undesirable, since it merely interferes with drainage.

If mutual adsorption does not occur, one may obtain selective separation in foam of a particular material by altering the temperature. This effect is apparently due to an abrupt decrease in surface viscosity.

Eddy Currents in Foams During Fractionation. If eddy currents of bubbles are formed in the foam in Abribat's apparatus, even for a short time, this may be sufficient reason to ruin a separation previously obtained. Such disturbances are greater in the more fluid portions of the foam, which have relatively large amounts of liquid, and are more easily produced for larger cross-sections of the foam column and for higher rates of gas flow. Abribat avoids eddy currents by using a device[8] which allows for a convenient agitation of the column of foam. The rotation of a series of vertical coaxial stirrers tends to carry each bubble in a circular path in a horizontal plane, and, as the foam rises, the path becomes helical. This arrangement is claimed to have the further advantage of increasing the length of travel of individual foam bubbles, besides which it allows drainage of the foam in a shorter column than would be necessary in the absence of such stirring. With this arrangement, the various fractions are sharply separated in the column.

The eddy currents are due to the various-sized bubbles formed with the porous cup. Large bubbles travel faster than small ones and collect at the top of the foam column. We have found that foams of uniform bubble size are formed by using a single orifice, or "spinnerettes." Such foams should be a distinct advantage in foam fractionations, since difficulties due to eddy currents would be largely eliminated.

Applications:

PROTEINS. For proteins, the rigidity of the interfacial films of the foams is a function of both pH and temperature. Other factors being equal, the most suitable conditions for foam fractionation obtain at the isoelectric pH. However, the temperature should not exceed a certain value above which some proteins such as gelatin do not form viscous films and the foam persistence is low. Furthermore, foam enrichment by foam drainage is almost nil. In such cases, above a certain temperature, separation is not improved at any concentration.

Gelatin has been separated by Abribat from heat-denaturable proteins, such as albumins and globulins. These proteins constitute the products obtained at the beginning of the fractionation, and some of them flocculate irreversibly in the foams. Gelatin is not denatured by heating its solution, nor is it denatured by foam formation.

COLLARGOL. For a sol of collargol, during the foam formation, the liquid is rapidly decreased in concentration of colloidal silver. No flocculation is produced in the liquid during the operation and until the removal of the last traces of silver. If proteins, which can be denatured by foaming, are present, the constituents of the foam may flocculate irreversibly.

METHYL VIOLET DYES. Abribat repeated the experiments of Kenrick[11] with a mixture of methyl violet (tri-, tetra-, penta- and hexamethyl-pararosanilines). The top fraction of the foam was almost entirely composed of the hexamethyl derivative (crystal violet) which had a characteristic violet color, whereas the product at the lower section of the foam was red-purple.

Towards the end of the operation, the fractionation could be accelerated and considerably improved by adding a very small amount of gelatin, which increased the volume of foam. Beyond mentioning the difference in color, Abribat does not indicate the purity of the fractions obtained.

MIXTURE OF SAPONIN, GELATIN, AND SODIUM OLEATE. When a dilute solution of gelatin containing sodium oleate was subjected to a foam-forming treatment, the top fractions of the foam column were composed exclusively of gelatin. The oleate was separated, subsequently, from an alkaline solution to avoid hydrolysis of the sodium oleate. In order to obtain suitable properties of the gelatin films in the foam, it was necessary to add about ½ cc. per liter of ethanol to the solution.

During the foaming of a solution of gelatin and saponin, the saponin was found in the top fraction. If the solution also contained sodium oleate, the dissolved substances were at least partly concentrated in the foam column in the following order: saponin, gelatin, sodium oleate. This appears to be in accordance with the rigidities of the interfacial films formed. According to Abribat, this does not follow the individual surface activities of the different substances, when these are taken separately.

OTHER MATERIALS. Abribat separated fatty acids closely related to each other in constitution and number of carbon atoms. He suggests the use of the method for various protein and lipoid substances, blood serum, various plasma, as well as for the separation of enzymes, virus, or hormones.

---

[11] Kenrick, F. B. J. Phys. Chem. 16: 513-17. 1912.

SCHÜTZ'S METHODS FOR FRACTIONATION OF MIXTURES IN FOAM

Procedure. F. Schütz[9, 10] and co-workers report on the fractionation by adsorption and crystallization on foam. Their method consists in bubbling nitrogen through a single orifice at the bottom of a relatively short column, 28 cms. high and 3.5 cms. in diameter (FIGURE 4), and the foam is allowed to go over a bend at the top of the tube into a side vessel. The height of the liquid through which each bubble rises, and the time the bubble rests in contact with the liquid, have to be sufficient for surface equilibrium to be approached. Schütz considers this condition satisfied with the liquid level about half-way up, but he does not give the basis for his statement.

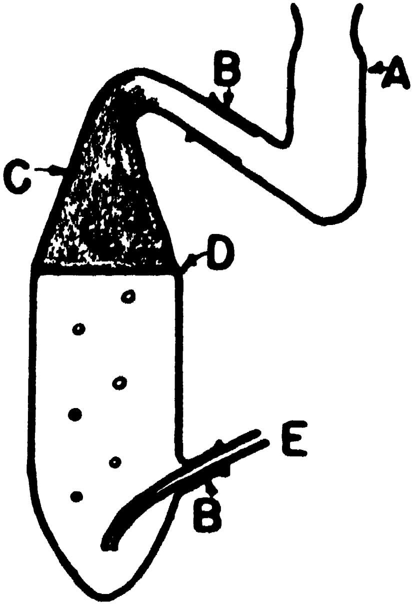


FIGURE 4. Foam fractionation apparatus, after Schütz[9].

The rate of gas flow is adjusted so that the bubbles reach the upper overflow bend in the tube. Schütz claims that this has the advantage of foam automatically ceasing to be carried over if its stability suddenly decreases or increases, when the surface active material collected from the foam has all been removed from the solution. At this point, the receiver is changed, the gas pressure is increased, and the next fraction collected. This accounts only for the volume of foam produced, regardless of bubble sizes and other properties, such as drainage characteristics.

If stable foam is not obtained from a solution, stability can sometimes be induced by adding a liquid which reduces the solubility of the substance to be adsorbed. Schütz finds that aqueous solutions of inulin hardly form foams, but they give very stable foams if prepared with 1.65 per cent ethanol or 10 per cent dioxane.

He notes that, unless there are indications from "peaks" of "foam time"* curves,[12] it is best to make the foam at as low a concentration as possible. This corresponds with the observations made by Abribat with regard to concentration.

In nearly all cases, some solid matter is found in the foam fraction, and this often leads to a more efficient fractionation. However, in foaming proteins, much of the solid matter in the foam fractions is surface-denatured protein. The foam can be broken by centrifuging or by drying in a desiccator.

Schütz recommends the choice of "peaks" of "foam time" against concentration curves[12] as a guide to the concentrations to be used for foaming. For the purification of proteins and bile salts, better fractionation was obtained at the isoelectric pH.

The surface tension of the foam fractions was lower than that of the original solution, and efficient fractionation was shown by an increase in the surface tension of the solution. However, Schütz considers that "foam time" is a more sensitive indicator than surface tension of when to change the receiver. On the other hand, we have seen before how in our work[4, 5, 6] the minima in surface tension concentration curves correspond to optimum selective fractionation of certain components from which the foam was formed.

Effect of Temperature. Thermolabile substances like enzyme solutions can be converted into foam at 0°C. For this purpose, the incoming gas should be pre-cooled. Changes in both bulk and surface properties may change considerably with temperature, as indicated, for example, in the drainage characteristics of foams.

The rates of flow of liquid through foams decrease with increase in both surface and bulk viscosities,[13] both of which are affected by changes in temperature.

If an increase in temperature decreases the viscosity without causing other pronounced changes in properties, it is advisable to use higher temperatures. However, Schütz finds that for some materials there is an optimum temperature, different for various materials, beyond which, at higher temperatures, fractionation by foam is no longer possible. For sodium cholate, this temperature is 36°C., and for saponin, 70°C. In this connection, it may be of interest to note that differences in drainage rates of foams for various materials become insignificant above certain temperatures, due to an abrupt decrease in surface viscosity.[13]

---

* The first method given by Schütz consisted in bubbling gas through the solution by means of a capillary tube. The period between the stoppage of gas flow and the time when half the foam-covered surface became free of bubbles was called the "foam time."

For aliphatic alcohols, Schütz says that a standardized shaking method was used. After shaking was stopped, the time taken until the center of the liquid became free from bubbles was called the "foam time."

[12] Schütz, F. Trans. Far Soc. 38: 85. 1942; 38: 94. 1942.

[13] Miles, G. D., L. Shedlovsky, & J. Ross. J. Phys. Chem. 49: 93-107. 1945.

Size of Bubbles and Rate of Drainage. The size of the bubbles can be varied by using jets with different openings. As a rule, smaller bubbles are more advantageous since they correspond to more surface for the same total gas volume. However, Schütz found that, for saponin, smaller bubbles remove less material than larger ones. This was probably due to the much slower drainage rate for the smaller-size bubbles.[13]

Rate of Foam Formation. Since Schütz used "foam time"[12] as a guide in the technique and this depends on the bubble size, sintered glass was found unsuitable, as it produces various bubble sizes which mask the observed changes in foam stability, quite apart from the fact that slow foaming cannot easily be achieved by means of such a filter. Somewhat better results can be obtained by using a multijet, with 12-20 openings of approximately equal diameter. In using more than one opening, the size of the apparatus must be increased, for, otherwise, the greater amount of foam produced per unit time reduces the desired preferential adsorption of one constituent. In our work,[13] stainless steel "spinnerettes" of the type used in making rayon (for example, 40 holes, 320×10^{-5} inches in diameter $\pm 5 \times 10^{-5}$ inches) are used for obtaining foams of uniform bubble size.

By forming foam from a solution of sodium tauro-glyco-cholate by a single jet, crystallization of fatty acids and bile acids in the foam fractions was obtained by Schütz, but, if a multijet was used for the same solution in this apparatus, none of the above results could be obtained. This was due mainly to the increased rate of flow of gas, and to the shorter time the foam remained on the surface of the liquid and in the solution vessel. The average time one bubble took to be drawn over into the side vessel was approximately 30-40 sec. If the rate of flow was increased so that only 5 seconds were allowed, there was hardly any difference in concentration between the foam fraction and the residual solution. When a multijet was used, the height but not the width of the apparatus was increased.

Schütz sometimes found that much of the foam became stagnant in the upper part of the vessel and that newly formed foam was moving through a channel of this stagnant foam. This was often found to be a disadvantage, because the stagnant foam which should have been removed with an earlier foam fraction was sometimes quite efficient in breaking the later foam bubbles when these came in contact with it. The stagnant foam was therefore removed through an opening at the top of the column (not shown in FIGURE 4), before it was carried over the side. With solutions which initially showed very low "foaming stability", it was found better to fill the vessel to a higher level, so that bubbles had a shorter distance to travel from the surface of the liquid until they reached the side vessel.

Applications:

BILE SALTS. R. Bader and F. Schütz[10] describe the fractionation by adsorption and crystallization on foam of bile salts. Free bile acid was adsorbed and found in crystalline form in the foam fractions. Foam was formed from 0.4 per cent sodium cholate at 24°C. with a jet 2-3 mm. in diameter.

Impure preparations yielded crystalline fatty acids in the earlier fractions and crystalline bile acids in the later foam fractions. When the fatty acids were removed by foaming and the bile acids started to separate, there was a sudden change in "foam stability". The nitrogen pressure for the first foam fraction was less than 0.5 lbs. per square inch and was adjusted until the flow of foam entering the side vessel was not quite continuous. The lower the rate of gas flow, the greater the fractionation. The concentration of the cholic acid in the foam fractions was 3 to 6 times that of the original solution. In subsequent operations, the pressure was increased or decreased for the next fraction, depending on whether the foam volume decreased or increased.

The optimal pH for these fractionations of salts of bile acids was 6.5. Intermittent addition of hydrochloric acid solution was found to be a more satisfactory method of controlling the pH than the use of buffer solutions, because the latter completely hindered the crystallization of the free acids in the foam fractions. In other cases, the salt in the buffer solutions may alter the adsorption in the foam. The fatty acids and similar materials appeared in the first three fractions.

The foam in the side vessels was mostly broken on standing 2 to 3 hours, and the remainder was broken by centrifugation.

DOGNON'S METHODS FOR FRACTIONATION OF CHEMICAL AND BACTERIOLOGICAL MIXTURES

A. Dognon[14, 15] discusses concentration and fractionation by forming foams for both chemical and bacteriological mixtures.

Procedure. In the simplest case, his apparatus consisted of a cylindrical tube 25 to 30 mm. in diameter and 50 to 60 cm. high. Drainage of foam was practically complete in the top section, if the gas pressure was regulated so that the upper boundary of the foam remained stable for a sufficiently long time. Then, the various fractions could be recovered by increasing the pressure so as to collect the successive foam fractions by overflow.

Applications:

BLOOD SERUM. Dognon confirmed the common observation that there was no significant separation of material in the foam above certain concentrations of solution. At higher concentrations, the adsorption was

---

[14] Dognon, A. Actualités Scientifiques et Industrielles. Hermann et C^{ie}. Paris. 932: 157-172. 1942; Bull. Soc. Chim. Biol. 23: 249-63. 1941; Rev. Sci. 79: 613-19. 1941.

[15] Dognon, A., & H. Dumontet. C. R. Soc. Biol. 135: 884. 1941.

poor, as indicated by Gibbs' Equation, and the drainage characteristics of the foam were unfavorable for effective separations. On diluting a solution, there was an increase not only in the ratio of the foam to solution concentration but also in the total concentration in the foam.

For blood serum, concentrations above 0.5 per cent cannot be treated. The adsorption of protein in the foam may give a concentration 20 or 30 times greater than in the initial serum. Considerable denaturation takes place, especially of the globulins.

BACTERIAL SUSPENSIONS. A separation of bacterial suspensions by foam is regarded by Dognon as a sort of microflotation. For this purpose, a cylindrical tube, shaped on top like an inverted funnel, had a two-hole rubber stopper at the bottom which contained a porous cup. Samples of liquid were periodically removed from below and examined for optical density with a photoelectric turbidimeter.

The bacteria removed from the surface of a solid medium were made into a suspension in the medium chosen. Centrifugation permitted the removal of agglomerated portions, whereupon the change in the suspension was determined by the turbidity of a diluted portion. It was not necessary to add foaming agents, for even in the distilled water there was sufficient foam to observe the effects of entrainment.

Koch bacilli (dried B.C.G.) suspended in water were easily concentrated in the foam and appeared in clumps which adhered to the wall of the tube. The remaining suspension rapidly became more dilute.

Under the same conditions, *B. coli, Staphylococcus alba,* and yeast cells (*Schizosaccharomyces*) were separated in the foam to only a very slight degree. However, by adding NaCl, Na_2SO_4, $CaCl_2$, or saponin, mucin or globulins, the separation was greatly enhanced. These effects increased at higher salt concentrations for all the above cases, as well as for Koch bacilli.

On the basis of the above experiments, it was possible to make a nearly complete separation of Koch bacilli from a mixed suspension containing one of the other organisms mentioned previously. A microscopic examination showed no Koch bacilli left in the residual liquid.

Dognon remarks that the separation of Koch bacilli from the other microorganisms in distilled water suspensions appears to be in accord with the concept of microflotation, indicating that the poorly wetted Koch bacilli are entrained in the foam. This may be attributed to the hydrophobic nature of the surface-covering of these bacilli. There is no simple explanation for the effect of electrolytes.

DISCUSSION

The methods for fractionation in foam which have been summarized suggest procedures which may be useful, but it is evident that no single method is applicable to all cases.

Abribat stresses the importance of taking advantage of differences in foam drainage. On the other hand, in Schütz's apparatus, the foam was removed as soon as possible after formation, so that only a short time was allowed for drainage. Schütz uses "foam time" and the foam volume produced at a fixed gas pressure as a guide to the procedure for fractionation of mixtures of materials in the foam. It appears that "foam time" is affected by drainage rates as well as by surface adsorption, but not either one alone.

It is advisable to separate foam which had been allowed to drain, as much as possible. It may be of interest to speculate with regard to the advantage in foam fractionations of foams with viscous films which drain slowly to a state where relatively large amounts of liquid are retained. Abribat says that, for protein foams, the films are, in effect, gels which are relatively rich in water, whereas, for solid interfacial films, the foam is nearly completely dehydrated when a steady state is reached. In such a case, solid is deposited on the tube or else a rigid fine spongelike structure is formed. It was noted previously that, in our work, a repetition of the operation of pouring the solution through the foam gives improved selective adsorption and a more complete approach to a steady state. All the material which can be removed from solution is not adsorbed in foam immediately. Consequently, slow-draining foams would be expected to lead to better fractionation in Abribat's apparatus, since it is obvious that this results in larger amounts of liquid remaining in the foam (per cm.^{2}) while it is drained, as well as after drainage has practically stopped. In this connection, Schütz made the observation that, the lower the rate of gas flow, the greater the fractionation, due to the longer time allowed for adsorption to reach a steady state.

Abribat hardly mentions adsorption, except to say that it does not determine the fractionation which, he states, is predominantly a function of surface "rigidities." Relative adsorption should be estimated on the basis of the same surface areas, so that cognizance should be taken of the volume of foam and the bubble sizes. In foam fractionation, the total adsorption will, of course, depend on the total surface. In FIGURE 3, taken from Abribat, the curves do not show differences in drainage rates on the basis of comparable total volume of foam, which is proportional to $H + h$, and, furthermore, the various fractions probably correspond to different bubble sizes. These factors, as well as bulk and surface viscosities, affect drainage rates,[13] so that it does not seem valid to attribute the overall differences in drainage rates, indicated by Abribat in FIGURE 3, to differences in "rigidities" of surface films, unless other data are available to show this.

A solution of a mixture of sodium dodecyl sulfate and sodium hexadecyl sulfate does not show rigid films of the type suggested by Abribat. There is no evidence of high surface viscosity in films made from the solution, but a purification of this mixture can be carried out by foam fractionation.[5] In this case, the solution was allowed to flow through the foam several times. The guide to effective separation was obtained on the basis of the surface tension-concentration curve, as described previously.

Gibbs' Equation deals with reversible adsorption, but there is reason to believe that some of the foam fractionations which Abribat and Schütz describe deal with irreversible adsorption of the type due to denaturation of protein.

It would be useful to have more examples of separations of mixtures of relatively pure materials of known composition, in order to learn more with regard to the principles involved. The suggestions given in the papers discussed should be pursued, so that the significant factors leading to clear-cut separations may be more precisely defined. In separations and purifications, it would be desirable to use better criteria of the purity of the fractions obtained.

It is clear that the methods of foam fractionation, purification with emulsions, and microflotation of bacteria can lead to many applications of theoretical and practical importance.

SUMMARY

Examples of the extent of adsorption at liquid-liquid and air-liquid interfaces have been noted from the literature.

Some of the possibilities of fractionations with foams and emulsions have been reviewed with particular reference to the more recent publications.

The various procedures are described and a number of factors involved are discussed, including selective adsorption, composition and concentration of solution, pH of solution, temperature, size of bubbles, rate of foam formation, eddy currents in inhomogeneous foams, and the effects of bulk and surface viscosity.

The surface tension-concentration and interfacial tension-concentration curves have been shown by Miles and Shedlovsky[4-6] to be a guide to the method for the purification of certain mixtures. Such data are also useful criteria of the purity of some materials.

Abribat[7] and Dognon[11] emphasize that foam drainage is a predominant factor in fractionation. They point out that drainage of foam represents a chromatographic column in which the constituents are stratified according to surface rigidity of the films rather than order of surface activity.

Their drainage data for foams cannot be clearly attributed to relative surface "rigidities", since the variations in surface areas and bubble sizes involved are not taken into account.

Transfer of all the material which can be adsorbed on foam from solution is not instantaneous, and it is suggested that low rates of gas flow and slow-draining foams lead to better fractionations, due to the longer time that foam bubbles are surrounded by adequate amounts of solution.

In the procedure used by Abribat and Dognon, it is claimed that eddy currents of bubbles formed in the foam may be sufficient reason to ruin a separation previously obtained. Abribat describes a device for the agitation of the column of foam which he uses to avoid eddy currents.

It is suggested that the difficulties with regard to eddy currents may be attributed to the formation of foams through fritted diaphragms which do not result in uniform bubble sizes.

Schütz uses "foam time" and foam volume produced at a fixed gas pressure as guides to the procedure for the foam fractionation of mixtures of materials.

Applications are given of foam fractionation of sodium alcohol sulfates, soaps, proteins, collargol, dyes, and bile acids.

The separation, by Dognon, of Koch bacilli from *B. coli*, *Staphylococcus alba*, and yeast cells (*Schizosaccharomyces*) by microflotation in foam is outlined.

The writer wishes to express his appreciation to Mr. Gilbert D. Miles and Dr. John Ross for their helpful suggestions in connection with this review.

ION-EXCHANGE ADSORBENTS AS LABORATORY TOOLS

By NORMAN APPLEZWEIG

*Laboratory of Cellular Physiology, New York University, New York, N. Y.**

Ion-exchange adsorbents at present enjoy a wide application in industry. The patent files are rich with examples of production tasks which ion-exchange has ingeniously been made to do. The use of exchange adsorbents as common laboratory tools, however, has yet to receive adequate recognition.

In order to illustrate the broad possibilities for these techniques, it was thought useful to review the principles involved in some of the commercial uses of ion-exchange. An attempt will also be made to discuss typical applications of ion-exchange in organic and biochemical research, in order to make these results known to a wider audience.

Since the concentration, isolation, and purification of organic materials from natural sources or dilute reaction mixtures are prime problems in organic and biochemical research, emphasis will be placed upon the applications of ion-exchange techniques to the accomplishment of these purposes. It will also be seen that ion-exchange provides simple means for the preparation of salts by double decomposition, removal of acid from acid salts, and for the separation of many substances.

While all the techniques reviewed here do not explicitly involve chromatography, it will be seen that they could, in most cases, be applied in the chromatographic manner.

The history of the development of ion-exchangers from the natural clays and humic materials through the synthetic siliceous, the carbonaceous and, finally, the synthetic resin exchangers is recorded in a review by Myers, Eastes and Myers.[63] The theoretical aspects of ion-exchange are discussed by Griessbach[40] and also by Myers in his chapter on synthetic resins in *Advances in Colloid Chemistry.*[61] A more recent discussion is to be found in the paper by L. E. Davis.[28] Studies on rates of exchange have been made by Walton,[86] by Nachod and Wood,[65] and by Thomas.[81] An excellent discussion of the theoretical conditions for exchange reactions is contained in Cannan's paper on "Chromatographic and Ion-Exchange Methods in Amino Acid Analysis".[22] For more complete bibliographies, reference should be made to the general reviews of Block[17] and Deitz.[29]

For the purpose of this discussion, we may consider ion-exchangers to be insoluble acids or bases. As such, they are capable of the reactions of ordinary acids or bases, with the notable exception that one of the reaction products always "leaves the solution" in the form of an insoluble salt. The salt-forming capacities of these insoluble acids and bases are, to a certain extent, a function of the dissociation constants of the reacting electrolytes. Simple exchange reactions may, therefore, be expected to behave in accordance with the mass action law.

---

* Present address: Hygrade Laboratories, Inc., New York, N. Y.

Sussman, Nachod, and Wood,[80] using the case of basic resins as an example, have pointed out that "it is rather difficult to conceive of an ionic system in which the anion is freely mobile in the usual sense while the cation comprises a substantially immobile charged locality on a macromolecular insoluble resin. It is possible that, under these conditions, ionization can consist of little more than a limited mobility of the anions among different cationic groups on the resin."

Since ion-exchange reactions involve a two-phase system, results comparable to those in physical adsorption or extraction by immiscible solvents are obtained. Thus, the rate of a given reaction is influenced as in physical adsorption, by time of contact, surface area, etc. Myers[61] suggested that, in order to determine whether a reaction is one of physical adsorption or of ion-exchange, it is necessary to analyze for all the constituents originally present as well as those which may be formed.

As in the case of "conventional adsorbents", ion-exchangers may be added directly to the solution, and the exchange allowed to reach equilibrium, or they may be used in columns, in which case continuous counter-current effects may be achieved.

In cation exchangers, the exchangeable hydrogen ion may be provided by any one of the following functional groups:

$$—SO_3H,\ —CH_2SO_3H,\ —COOH,\ —OH,\ —CH_2OH, \text{ or } —CH_2SH.$$

The acid-binding properties of anion exchangers may be due to aromatic or aliphatic basic nitrogen groups such as

$$—NH_2,\ =NH \text{ or } \equiv N.$$

In the ion-exchange reactions listed below, *A* represents an insoluble acidic cation exchanger, and *B* an insoluble basic anion exchanger.

Base exchange

$$Na \cdot A + CaCl_2 \rightarrow Ca \cdot A + NaCl \qquad (1)$$

Acid formation

$$H \cdot A + NaCl \rightarrow Na \cdot A + HCl \qquad (2)$$

Acid removal

$$B + HCl \rightarrow B \cdot HCl \qquad (3)$$

Anion exchange

$$B \cdot HCl + NaI \rightarrow B \cdot HI + NaCl \qquad (4)$$

The remaining discussion will be organized mainly around these four reactions.

BASE EXCHANGE

Reaction 1 represents the classical use of zeolite exchangers for the softening of water. In like manner, calcium ions in milk or blood may be exchanged for sodium. In the case of milk, a soft curd milk was produced which was found useful in infant feeding.[55] The treatment of blood with a sodium exchanger was suggested by Steinberg[77] as a means for obtaining "uncoagulated blood."

The quantitative aspects of this exchange have been investigated by Applezweig and Rice,[5] who found that approximately 80 per cent of the blood calcium is removed. Since the amount of calcium which remains is theoretically sufficient to promote clotting, it is assumed that this residual calcium is in a bound or non-ionized form and, as such, not capable of promoting coagulation.

In reaching conclusions concerning the concentration of calcium ion required for the clotting process after the addition of oxalates or citrates, it was necessary to assume certain relationships between the two systems involved, *i.e.*, calcium oxalate or citrate on the one hand and calcium proteinate on the other. Ion-exchange offers a new approach to the study of this mechanism, since calcium is removed simply by exchange with sodium without addition of extraneous anions. Thus, only the natural equilibria in the blood need be considered.

In a like manner, the role of various cations in other physiological phenomena may be explored. It has been generally assumed, for example, that serum calcium is associated with complement function. Ecker and Pillemer[33] studied the effects of calcium "disionizing" anions upon complement and reasoned that calcium plays a role of doubtful significance in complement fixation. In Dr. Ecker's laboratory, human, rabbit, and guinea pig sera which had been treated with a sodium exchanger were found to have lost none of their complement activity.[34] This would seem to confirm the belief that ionic Ca^{++} plays a negligible role. Breazeale, Pierce, and Breazeale,[18] in studies on serological precipitation-flocculation tests, found that the reactivity of sera could be shifted from negative to positive and *vice versa* by treatment with zeolites. These investigators found that Kahn antigen adsorbs "reagin" out of syphilitic serum, thereby producing a seronegative serum. By treating this experimentally produced sero-negative serum with calcium zeolite, they again obtained positive flocculation reactions to the various tests. Later, they were able to apply this same principle to a test for syphilis by employing a suspension of pure sodium zeolite as an indicator.

Pierce and Breazeale[67] then investigated the removal of divalent cations from solutions by beef heart antigens and found Kahn and Hinton antigens to be capable of removing calcium and magnesium from aqueous solution. The floc which is formed by this reaction releases these divalent ions again when it is washed with saline. On the basis of these observations, a new theory of the mechanism of syphilis serology was advanced, namely, that, in positive sera, there must be either an increase in ionization of divalent cations or some profound vectorial shift with respect to the ionizatiŏn constant of calcium compounds causing the antigen to floc. At any rate, it was shown that beef heart antigens act purely in the capacity of an ion-exchanger and shift from a sodium salt to either a calcium and/or magnesium salt, thereby producing a floc.

An obvious application for base exchange is in the preparation of salts by double decomposition. Thus, the sodium salt of penicillin may be converted to the calcium salt by passage over a calcium exchanger. In some cases, it may be desirable to convert a soluble salt into a less soluble one. This is the principle upon which many of the chromatographic separations using ion-exchange adsorbents are based. The formation of a less soluble salt delays the passage of the compound through the column.

When cations are present as contaminants in crude reaction mixtures, base exchange can be especially useful. For example, the various cationic contaminants may be exchanged for a single one capable of easy removal. Harmful or poisonous cations may be exchanged for harmless ones.

Sugar juices have been commercially purified by means of base-exchange. A natural Ca zeolite was used by Harm, in 1896,[41] to replace molasses-forming sodium and potassium ions in sugar solutions with calcium ions. Jeanprost[49] reported on the treatment of diluted molasses with a Pb-zeolite. Lead salts in solution were later removed by passing the liquid over an ammonium exchanger. Austerweil[9, 10] used NH_4-zeolites to reduce substantially the ash content of sugar solutions. NH_4 ions which passed into the sugar juice were removed in the form of ammonia by evaporation.

Jones[50] has found that the conversion of calcium and magnesium chlorides in petroleum to sodium chloride by a sodium exchanger results in reduction of corrosion in refinery equipment. Bahrdt[12] has used cation exchange for the removal of interfering metals in an analytical determination of sulfate.

One of the earliest adaptations of base exchange to laboratory problems was that of Folin,[37] who, in 1917, developed a method for the determination of ammonia in urine. Sodium permutit (later called Folin Decalso), is used to "adsorb" the ammonia. After washing, the adsorbent is treated with sodium hydroxide and Nessler's reagent, the ammonia nitrogen being determined colorimetrically.

Ammonia in other fluids may also be determined in this manner. A common way to prepare ammonia-free water is by adding a powdered sodium zeolite.

Bird[15] has used cation exchangers in the preparation of sols of silicic acid and other acidic oxides. According to his patent, relatively pure silica sols are obtained by treatment of a dilute solution of sodium silicate with a hydrogen exchanger. The sodium ion in the silicate solution is replaced by hydrogen to give a silicic acid sol in the effluent:

$$H \bullet A + Na_2SiO_3 \rightarrow Na \bullet A + H_2SiO_3$$

The introduction of carbonaceous and synthetic resinous zeolites capable of hydrogen exchange vastly broadened the possible fields of application for ion-exchangers. These newer-type adsorbents are available as relatively pure compounds possessing reproducible physical and chemical properties. Like the siliceous zeolites, they may be used in sodium or calcium cycles as base exchangers with the advantage that they have a greater exchange capacity per unit volume and are stable over wide ranges of pH.

When used in the hydrogen cycle as shown in REACTION 2, acids are formed. This reaction is especially important when employed together with acid-removing adsorbents (REACTION 3) to effect complete removal of electrolytes from solution. This will be discussed, later, in connection with demineralization.

The exchange of metallic cations for hydrogen offers many other interesting possibilities.

Foremost among these is the possibility of an "acid machine". If a solution containing cations is passed through a hydrogen exchange column, the effluent will be found to contain an amount of acid equivalent to the cations originally present.

Polis and Reinhold[69] have used this principle to determine total base in serum. 0.2 ml. of serum is added dropwise to a micro column of a synthetic hydrogen exchange resin. The column is washed with water and the effluent titrated with standard alkali. Since base which is combined with carbonate or bicarbonate will yield H_2CO_3, it is necessary to do a CO_2 combining-power determination on another portion of the same serum. The two results, when taken together, represent a value for total base which compares very favorably with that obtained by the more cumbersome method of electrodialysis.

A hydrogen exchanger is made to serve a dual purpose in the manufacture of high-grade pectin.[60] Metallic cation impurities are removed and acid for hydrolysis is provided by the hydrogen exchange. Similarly, the production of a palatable artichoke syrup[35] and the inversion of sucrose solutions may be aided by hydrogen exchangers.

The use of zeolites in the concentration of metal traces offers promise. Little work has been done in this direction. Abrahamczik[1] has shown that by permutitic exchange minute amounts of iron can be isolated and determined. Ten liters of distilled water containing 0.0025 parts per million of ferrous iron were run through a 50 x 450 mm. column of "neopermutit" (about 400 g.). One hundred ml. of warm saturated sodium chloride solution were then passed slowly through the column in the reverse direction, followed by 150 ml. of wash water. Iron was determined colorimetrically in an aliquot of the combined solutions, and 90 per cent of that added was found.

ADSORPTION OF ORGANIC CATIONS

An observation, in 1921, that adrenaline could be removed from solution by shaking with permutit prompted Whitehorn to investigate the use of this base exchanger as a reagent for amines.[88] He investigated not only what kinds of substances can be removed from solution by base exchange, but also in what ways this reaction can be affected by chemical and physical conditions. The adsorption of fifty-two organic compounds was studied qualitatively, and fundamental facts concerning the quantitative aspects of base exchange were disclosed.

Whitehorn showed that permutit could be used to separate relatively strong nitrogenous bases, having a dissociation constant of 5 x 10^{-9} or greater, from weaker bases and non-basic substances. He found that more of a given base could be removed from neutral solution than from acid or alkaline solution. Bases could also be removed from solution in varying concentrations of ethyl alcohol, in amyl alcohol, and in ether.

For the quantitative removal of bases from solution, it was shown that filtration through the exchanger is preferable to agitation because (1) there is no accumulation of the sodium salt formed by the reaction which would tend to reverse it; (2) successive portions of a relatively large volume of fluid can be rapidly brought into intimate contact with the exchanger; and (3) easily oxidizable substances are saved from destruction.

To liberate the bases from the exchanger, Whitehorn investigated the use of sodium hydroxide, sodium cyanide, and potassium chloride. Of these, he found a saturated aqueous solution of potassium chloride to be the most generally useful reagent. Using this technique, he reported the recovery of 92 per cent of adsorbed ethylamine nitrogen and 85 per cent of adsorbed choline nitrogen.

Later,[90] Whitehorn developed a chemical method for estimating epinephrine in blood using silicic acid as the exchange adsorbent and sulfuric acid as the eluent.

The fundamental observations of Whitehorn provide the necessary background for the use of ion-exchangers in the recovery of organic cations.

One of the most useful applications of this technique has been in the isolation of alkaloids. A typical application is that of Oberst, who used permutit in the determination of morphine in the urine of morphine addicts.[66] Saturated sodium carbonate was used to liberate the alkaloid in the presence of the Folin-Denis phenol reagent. The blue color formed by the reaction of the alkaloid with the reagent was read colorimetrically.

The commercial application of ion-exchange to alkaloid recovery was greatly enhanced by the development of the synthetic hydrogen exchangers. Relatively large amounts of alkaloids could be adsorbed from dilute acidic solutions. Higgins[46] and Riley[72] patented the use of nicotine-containing hydrogen exchangers as horticultural poisons. Tiger and Dean[82] disclosed a procedure for obtaining concentrated solutions of nicotine salts by passage of an aqueous extract of tobacco over a hydrogen exchanger and subsequently liberating the nicotine in the form of its hydrochloride by means of constant boiling hydrochloric acid.

The use of a hydrogen exchanger to adsorb quinine from acid solution and to concentrate the alkaloids in a totaquine preparation was reported by Applezweig.[4] The recovery of the alkaloids from the exchanger was accomplished by means of alkali regeneration and solution in an organic solvent.

This communication also suggested the application of ion-exchange to alkaloid extraction by means of a cyclic system. Such a technique would permit the acid percolate to pass through a column of exchanger and return to the percolator in a continuous fashion. Thus, the drug would be constantly exposed to an extraction fluid free from alkaloids but saturated with respect to non-cationic ingredients.

Based upon these principles, a commercial installation for the isolation of scopolamine from cultivated datura plants was put into operation in the summer of 1944. The procedure proved sufficiently versatile to permit its adaptation to a field process for the manufacture of Cinchona alkaloids from low-quality green bark in jungle areas.[6] A portable unit was designed at the Engineer Board, Fort Belvoir, Va., which is capable of extracting the alkaloids contained in 13,000 lbs. of green bark per month with the use of approximately 1,000 lbs. of chemicals. The losses and much of the expense involved in the drying, grinding, and shipping of Cinchona bark to extraction plants may thus be eliminated.

The findings of Whitehorn that organic bases of moderate strength could be removed from solution by base exchange suggested to Cerecedo and Hennessy the idea of using zeolites for the isolation of vitamin B_1.[25] Extracts of rice polishings, brewer's yeast, and wheat germ were passed hot over a bed of Decalso (a synthetic siliceous zeolite) which had previously been stirred with sulfuric acid at a pH of 4. The adsorbed vitamin was then released by exchange with ammonium ion from a hot molar solution of NH_4NO_3. A single silver precipitation, followed by a precipitation with silicotungstic acid, yielded highly potent concentrates, from which, on recrystallization, crystals of pure vitamin hydrochloride could be obtained.

The successful commercial synthesis of thiamin precluded the wide use of base exchange for the recovery of this compound from natural sources, but the method of Cerecedo and Hennessy has become a standard one for the determination of vitamin B_1 in foodstuffs and physiological fluids.[43, 44]

Recently, Herr[45] studied the use of synthetic cation exchange resins in the separation, recovery, and concentration of thiamin. It was found that the acid-regenerated form of the resin completely removes thiamin, for which the resin has a large capacity, in the presence of riboflavin. Since riboflavin is not adsorbed under the conditions in which thiamin is completely removed, the separation is satisfactory. Thiamin is eluted in excellent yields by passage of strong mineral acid through the column.

The affinity of riboflavin for the resinous hydrogen exchanger was found to be much less than that of thiamin. Also, elution of riboflavin could be accomplished with a lower concentration of acid than that needed for thiamin recovery.

The purity of the eluate was found by Herr to be satisfactory for thiochrome analysis. For the recovery of thiamin from natural sources, the resin offers a distinct potential advantage over siliceous zeolite as an adsorbent; namely, elution can be accomplished with a volatile acid instead of a salt which cannot be removed readily from the eluate.

Although Binkley and co-workers have reported on the use of a Decalso column in the isolation of vitamin K,[14] in this case the adsorbent did not serve as an ion exchanger. The vitamin was removed from a petroleum ether solution and eluted by means of benzene in petroleum ether or acetone.

In 1932, Lejwa[58] purified the gonadotropic hormone by shaking urine with permutit for 2 hours and then eluting the adsorbed hormone with dilute NH_4OH. Active crystalline material was thus obtained which assayed 1,000 mouse units per mg.

Katzman *et al.*[51] were able to prepare concentrates of the hormone possessing 8,500 I.U. per mg. by adsorption on permutit and elution with an alcoholic solution of ammonium acetate. Adsorption from acidified pregnancy urine was found to take place at pH 4, but very little hormone was adsorbed at pH 5. If the hormone functions as a cation in the ionic exchange, then this would indicate that pH 4 is below its isoelectric point. Since evidence by other investigators places the isoelectric point between pH 3 and 3.5, the authors admit that it is possible that some other physical phenomenon is responsible for the adsorption.

Potts and Gallagher[70] used "Folin permutit" to separate, in high yield, the oxytocic and pressor principles of pituitary extract. The pressor hormone, a more basic ampholyte than the oxytocic principle, is adsorbed on the base exchanger. The separated principles offer a marked advantage for the further purification of the two products, since the losses are slight and the separation of the two activities is fairly complete. The pressor principle was adsorbed from aqueous solution and eluted with 10 per cent sodium chloride solution.

The adsorption of the active principle of hypertensin upon a hydrogen exchange resin has been announced by Cruz Coke, Gonzalez, and Hulsen.[27] No effective elution procedure was, however, found by these investigators.

A method for the determination of citrulline and allantoin based upon the cation exchange removal of the former from serum containing both compounds was described by Archibald, in 1944.[7] In developing this method, Archibald investigated quantitatively the adsorption by ZeoKarb and Amberlite IR-100 of citrulline, allantoin, allantoic acid, alloxanic acid, alloxan, alloxantin, parabanic acid, methylurea, phenylurea, and thymol.

McColloch and Kertesz[58] reported separation of pectin-methylesterase and pectin-polygalacturonase from a commercial pectinase by the action of a cation exchange resin. The authors explained their results in the following way:

The removal of pectin-methylesterase, but not of pectin-polygalacturonase, will be accomplished where the pH of the solution is below the isoelectric point of the methylesterase but still above that of the polygalacturonase. The authors had observed the different behavior of these two enzymes towards acids, and the fact that they had different pH optima made it likely that they would behave differently on ion exchange columns. They believe that the method will prove applicable to the separation of some other enzymes.

Among the bases which Whitehorn found removable from solution by permutit, were the basic amino acids, arginine, histidine, and lysine. Dubnoff[32] applied this finding to the separation of arginine from a protein hydrolysate for subsequent estimation by the Sakaguchi reaction.

Archibald was able to prepare glutamine solution free of contaminating arginine by passage of the contaminated solution through a bed of Decalso.[8] The effluent contained 90 per cent of the glutamine, but no arginine. The column was then washed with water to yield all of the glutamine left in the column and 25 per cent of the adsorbed arginine. A final washing with 3 per cent sodium chloride yielded the remaining 75 per cent of adsorbed arginine and no glutamine.

The use of strong neutral salt solution for the removal of the basic amino acids from siliceous sodium zeolites interferes seriously with subsequent isolation and purification steps. Block, in 1942, developed a method for the separation of basic amino acids from the other constituents of protein hydrolysates based on the use of hydrogen exchange resins.[16, 17] The basic amino acids are recovered in the form of a concentrate by treating the adsorbates with 7 per cent HCl.

At about this time, a number of workers were investigating the possibility of using ion exchange to sort amino acids into three groups: the acidic amino acids, held by the acid-binding resins; the basic amino acids by cation exchange; and the essentially neutral amino acids which might reasonably be expected not to be held by exchange adsorbents. Using synthetic resin exchangers, Freudenberg, Walch, and Molter[38] found that all amino acids could be taken up by a cation exchanger and preferential elution would have to be used in recovery of different fractions. Tryptophane was specifically separated from other amino acids by Turba, Richter, and Kuchar, using an organic cation exchanger.[85]

Englis and Fiess[36] and, later, Cleaver, Hardy, and Cassidy[26] reported on the conduct of amino acids in various exchange reactions. Their work confirmed the fact that hydrogen exchangers were capable of adsorbing all amino acids to some extent.

The latter investigators explain this seeming deviation from the hypothetical by the fact that the acidifying action at the surface of the resin brings all amino acids at least partially into the cation form and thus makes them "exchangeable". Also, "onium" formation may occur at the surface of the resin, converting a strongly acid resin, strongly acid due to sulfonic acid groups, into a less strongly acid adsorbate-resin, acid now due to carboxylic groups. The ammonium and sodium forms of the resins, on the other hand, were found by Cleaver, Hardy, and Cassidy to be quite selective for the cationic amino acids.

These authors also discussed their data with a view to the use of ion exchangers for chromatographic separations. They predict the possibility of separating arginine and histidine in this manner, since they have different adsorption isotherm slopes.

Sims[75] has improved on the Dubnoff determination of arginine by chromatographically separating it from glycocyamine. Both compounds give the Sakaguchi reaction. By careful control of the flow rate and salt concentration in a sodium resin exchange cycle, the glycocyamine is successively passed down the column and eluted in advance of the arginine.

ACID ADSORPTION AND DEMINERALIZATION

Since whole molecules of acid are adsorbed by the amine-type resins, it may be said that these exchangers really act as acid adsorbents (REACTION 3). Acids so adsorbed may be recovered by displacement with another acid, or in the form of a salt by regeneration of the exchanger with an alkali.

Unlike the base exchangers, there were no anion exchangers in common use as adsorbents prior to the advent of synthetic resinous exchangers.

The researches of Adams and Holmes, in 1935,[2] and of Griessbach, in 1939,[39] made possible the production of basic resinous adsorbents. Since these adsorbents have been commercially available, many possibilities for these valuable tools have been realized.

As has been mentioned previously, the use of acid-binding resins for the removal of acidic amino acids is an important one. When Block[16] isolated basic amino acids by means of a hydrogen exchange resin, he first pre-treated the protein hydrolysate to remove excess acid. This was accomplished by vacuum evaporation and finally by stirring with acid-binding resin until the reaction of the solution was approximately pH 6. In this same year, Freudenberg, Walch, and Molter[38] showed that acidic amino acids were adsorbed preferentially by an acid-binding resin. Cannan[21] used a basic resin for the separation of the dicarboxylic amino acids from protein hydrolysates. Estimations of the glutamic acid and aspartic acid in egg albumin, B-lactoglobulin, and edestin were reported. The adsorbed dicarboxylic acids are eluted, according to Cannan's procedure, with 0.25 M hydrochloric acid. The amino acids in the eluate are readily crystallizable, or they may be estimated by determination of amino nitrogen and carboxylic values, since the resin was found to be remarkably selective for the dicarboxylic acids.

The observations of Cannan were confirmed and the quantitative aspects of the anion exchange of amino acids studied by Englis and Fiess[36] and Cleaver, Hardy, and Cassidy.[20] The latter group concludes that much work remains to be done before resins can be used with complete confidence in the analytical separation of amino acids, but they offer many stimulating suggestions regarding new techniques which should be tried.

A neat illustration of one of the many ways in which acid-binding resins can be used is that of Buc, Ford, and Wise.[19] This group, while working upon an improved synthesis of β-alanine, was faced with the problem of liberating β-alanine from its hydrochloride. The use of lead oxide for this purpose was found to be tedious and troublesome. Attempts to use aniline or pyridine in methanol were unsuccessful. A method using pinene for this purpose was reported by Austin[11] to give only a 50 per cent yield. Buc, Ford, and Wise found that, when an aqueous solution of pure β-alanine hydrochloride is passed through a bed of the acid-binding resin, the resulting effluent is substantially chloride-free. The yield of chloride-free β-alanine is 93 per cent when pure β-alanine hydrochloride is used, or 83-88 per cent when the crude hydrochloride is used. Since the monoaminomonocarboxylic and the basic amino acids are not adsorbed by acid-binding resins, this method would prove generally useful for obtaining other amino acids from their anion salts.

When REACTION 3 (acid adsorption) is used in combination with REACTION 2 (hydrogen exchange) the technique referred to as *demineralization* or *deionization* becomes possible. Thus, when a salt solution is passed through a tandem setup, first through a hydrogen exchanger and then through an acid adsorbent, an acid is formed in column one where the cation has been exchanged for hydrogen, and the acid is removed in column two, thus removing the anion and yielding a solution free from electrolytes.

The possibilities for demineralization were immediately realized by the commercial water-conditioning organizations, whose engineering staffs have now made possible the production, at very low cost, of demineralized water which is said to compare favorably with that obtained previously by distillation.[42, 83]

Liebig, Vanselow, and Chapman[54] have pointed out the suitability of water purified by ion exchange resins for the caring of plants in controlled nutrient cultures. Objectionable traces of copper and other heavy metals which may occur in distilled water are absent from ion exchange water.

The removal of ionic contaminants from solutions of non-electrolytes should be of considerable interest to the organic chemist and biochemist. Smit[76] proposed the demineralization of solutions of non-electrolytes by passage over ion exchange adsorbents. His patent describes the purification of glucose solutions, paraffin, and other impure liquids, by treatment with a hydrogen exchanger and subsequently removing the free acid by means of an acid adsorbent.

Another application of ion exchange for the complete demineralization of non-electrolyte solutions was that of Holmes[48] who purified gelatin solutions by passage alternately over a bed of cation and anion exchange adsorbent.

Rawlings and Shafer, in 1942,[71] reported on the demineralization of clarified sugar cane juice. They found that, in a commercial process, cation and anion exchangers removed 95 per cent of the ash, 32 to 50 per cent of the original impurities, and 50 per cent of the color. A similar process for the purification of beet juice was described by Weitz, in 1943.[87] Applying such a procedure in the laboratory, Platt and Glock[68] used demineralization to purify an extract obtained from muscle preliminary to the estimation of the inositol content. By this treatment, the inositol-containing solution was obtained in a purer condition than by the conventional precipitation procedures previously used.

Williams and Johnson[90] have devised an improved technique for the determination of pectin. They found that pectin can be quantitatively determined by electrodeposition, provided the electrolyte concentration of the solution is low. Ion exchange resins were used to effect this lowering of electrolyte concentration. Here, again, is a method of broad possibilities: the isolation of electrically charged colloids by electrodeposition from solutions which have been "deashed" by ion exchange.

Tandem column operation is often useful where the desired ingredient of a mixture is exchangeable. McCready and Hassid[59] found that the preparation of the Cori ester from starch by Hanes' procedure can be very much shortened by using ion exchange adsorbents. Following phosphorolysis of the starch in the presence of phosphate buffer and the precipitation of inorganic phosphate, the reaction mixture was first passed through a column of hydrogen exchange adsorbent, which removed cations from solution. The effluent was then run through an acid-adsorbing column, where the glucose 1-phosphate, being a strong acid with high exchange capacity, was adsorbed. Soluble impurities including dextrins, proteins, and weak organic acids passed through unchanged. The glucose 1-phosphate was then eluted from the adsorbent with dilute alkali and isolated as the crystalline dipotassium salt.

Since cation exchangers had been used in the preparation of sols of acidic oxides,[15] Ryznar used anion exchangers for the preparation of hydrous oxide sols.[78] An alkali-regenerated anion exchanger was used in the preparation of sols of hydrous aluminum oxides, ferric oxides, and other metal oxides.

If the alkali-regenerated anion exchanger is considered to act only as an acid-binding agent, the reactions may be expressed as follows:

(1) $2AlCl_3 + 6H_2O \rightleftharpoons 6HCl + Al_2O_3 + 3H_2O$

(2) $6B + 6HCl \rightarrow 6B \cdot HCl$

SO_4^{--}, Cl^-, NO_3^-, and other anions may be removed in this way. The process is most successful when very dilute solutions are used. When concentrations of more than 1 per cent are used, aluminum and ferric sols show a tendency to precipitate and clog the bed.

Ryznar further demonstrated that a combination of cation and anion exchangers could be used to purify colloids formed by other methods. Colloidal sols formed by double decomposition or hydrolysis can be passed through the two exchange beds in series with a resultant removal of most of the electrolytes. This method was found to be simpler and quicker than the use of dialysis for the same purpose.

The preparation and properties of a protein-formaldehyde acid-binding resin were reported by Bhatnagar and associates.[18]

Adsorption experiments with fumaric, maleic, and citraconic acids showed that with NH_3-condensed and amino resins the adsorption of *cis* acid is greater than that of *trans* acid. This paper also discussed the adsorption of oxalic, malonic, succinic, aconitic, mono-, di- and trichloracetic, cyanacetic, hydroxyacetic, aminoacetic, propionic, chloropropionic, lactic, amino-propionic, pyruvic, levulinic, dibromosuccinic, malic, tartaric, aminosuccinic, and citric acids.

When it is desired to isolate organic acids which are in solution in the form of their salts, the demineralization procedure may also be resorted to. The recovery of tartrates from grape wastes by ion exchange was studied by Matchett *et al.*[57] Perhaps 10 million pounds per year of tartaric acid are potentially available from the waste pomace and still slops of the grape juice pressing and wine and brandy making industries. The tartaric acid, which is present mostly as potassium bitartrate, is first converted to the free acid by passage through a hydrogen exchanger. The acid is then adsorbed on an anion exchanger from which it may be released by regeneration with sodium carbonate. A solution of the sodium salt is obtained which is from ten to twenty times as concentrated as the original slop and substantially free from the many impurities that would accompany it through a similar evaporative concentration. In the final step, calcium chloride is added to precipitate the tartrate in the form of a relatively pure calcium salt.

ANION EXCHANGE

In their study of tartrate recovery, Matchett and his group made several interesting observations regarding anion-exchange phenomena. When solutions of potassium bitartrate were passed over anion exchangers which had been regenerated with sodium hydroxide, tartrate ion was fixed to an extent equivalent to precisely half of the total present, while the other half passed through as the normal salt. When a sodium carbonate regenerated anion exchanger was used, a larger amount of tartrate was removed and an equivalent quantity of carbonate ion appeared in the effluent.

From this, the authors infer that both regeneration and loading are definite anion-exchange phenomena. "Beds regenerated with carbonate retain an equilibrium load of carbonate ion, and this can be replaced under the conditions outlined by tartrate. Beds regenerated with caustic alkali solutions would not retain anions and acid adsorption would proceed through fixation of hydrogen ion with consequent retention of an equivalent anion."

Further experiments by this group have shown that it is possible to set up a definite replacement series of anions and that the course of exchange will be fixed by the relative positions of the ions involved and their activity in the solution. For example, tartrates may be recoverable from *unacidified* still slop by exchange with chloride on an anion exchanger regenerated by means of a sodium chloride solution.

In a similar study, Buck and Mottern[20] attempted the purification of apple syrup by ion exchange. It was hoped that ion exchange would remove malic acid and also reduce the lead and arsenic content. It was found that about 80 to 90 per cent of the acid is removed and a palatable syrup is produced by treatment with a single anion exchanger. A variable amount of arsenic is removed by the single treatment. To ensure a more complete removal when it becomes necessary, a two-step (cation-anion) or a three-step (anion-cation-anion) exchanger treatment must be used. As much as 96 per cent of the arsenic has been removed by the three-step method, and a smaller amount has been eliminated by the two-step method. In addition, the cation exchanger removes about 90 per cent of the ash.

Additional evidence for the use of basic resins as anion exchangers has been provided by the work of Sussman, Nachod, and Wood[80] on metal recovery by anion exchange. Metals which form complex ions may be recovered by true anion exchange. Use was made of the fact that treatment of an anion exchanger salt with a solution of a salt whose anion has a higher valence or weight than that of the anion already adsorbed on the anion exchanger, results in the displacement of the adsorbed anion by that of the salt in the influent solutions. Using chromium as an example, best results were obtained with adsorption from a neutral chromate solution on the exchanger chloride. Alkaline regeneration yielded a concentration of the complex ion in the effluent which was more than twenty times that in the influent.

Certainly a most important anion is that of penicillin. While it is known that the use of ion exchangers in the isolation and purification of penicillin has been thoroughly investigated, this information has not yet been made public. A short note in *Science,* in 1944, reported that a South American group had used ion exchange resins for the purification of crude culture filtrates containing penicillin.[27] The filtrate so treated retained all of its antibiotic activity but was free of toxicity. No details of value were, however, given.

Carter *et al.*,[21] in a paper on *Isolation and Purification of Streptomycin,* describe a chromatographic procedure which is essentially one of ion exchange. "Alkaline alumina removes streptomycin from neutral aqueous solutions, and elution with aqueous acid is slow and incomplete. Acid-washed alumina does not remove streptomycin from an aqueous solution but does so from aqueous methanol. This information provided the basis for a chromatographic method of purification. If a faintly acid solution of crude streptomycin chloride in 70 to 80 per cent methanol is percolated over a sulfuric acid-washed alumina column (pH 5 to 6), an inactive fraction first appears followed by an active fraction. The active material contains sulfate ion but no chloride. Evidently chloride ion has been replaced by sulfate from the column and the streptomycin sulfate thus formed then passes through the column less rapidly since the sulfate is much less soluble than the chloride in methanol."

ION EXCHANGE AND CHROMATOGRAPHIC TECHNIQUES

Because synthetic exchangers have a large capacity and react in a predictable fashion, their use in procedures such as the one noted above for streptomycin appears inviting. Thus, by exchange, less soluble salts may be formed which delay the passage of the compound through the column.

Another variation in technique is possible in the case of compounds such as the alkaloids. Here, adsorption takes place from aqueous solution by exchange with hydrogen ion in a cation exchanger. Elution, however, requires that the adsorbent be made alkaline and then leached with a solvent. The alkaloids may be separated by collection of portions of the eluate in the manner of the typical liquid chromatogram.

Direct visual observation of chromatographic banding on synthetic resins has been reported by Myers, Eastes, and Urquhart.[64] Light orange anion exchange resins were noticed to form a reddish-orange band superimposed on a brown band when a mixture of hydrochloric and sulfuric acids was flowed through the column of the resin. Analysis revealed that the reddish-orange band was the pure hydrosulfate, the brown band the pure hydrochloride form of the resin.

Even on darkly-colored cation exchange adsorbents, adsorption may be followed visually by the use of ultraviolet light. The adsorption of quinine upon sulfonated coal and its subsequent elution has thus been observed by the author.

The use of light-colored resins treated with indicators and of finely powdered exchangers in classical chromatographic techniques is an inviting field for discovery. Even more promising are the prospects for their

APPENDIX: LIST OF COMMERCIALLY AVAILABLE ION-EXCHANGE ADSORBENTS

Trade name	Manufacturer*	Type	Composition	Capacity,† gram milli-equivalents/ml. of exchanger
Amberlite IR-100	5	Synthetic resin cation exchanger	Modified phenol formaldehyde-sulfonic	0.44 @ 1.38 gm. m. eq. regenerant/ml. of exchanges
Amberlite IR-4B	5	Synthetic resin acid adsorbent or anion exchanger	Modified phenol formaldehyde polyamine condensate	1. HCl—1.21 @ 1.66 gm. m. eq. Na_2CO_3 as regenerant 2. H_2SO_4—1.54 @ 1.66 gm. m. eq. Na_2CO_3 as regenerant
De-Acidite	4	Synthetic resin acid adsorbent or anion exchanger	Aliphatic amine resin	0.55 to 1.2
Decalso	4	Synthetic sodium aluminum silicate cation exchanger	$Na_2O:Al_2O_3:6\ SiO_2$	0.40 to 0.55
Duolite A_1	2	Synthetic resin acid adsorbents or anion exchangers		1.68 to 1.8
A_2				2.0 to 2.2
A_3				1.4 to 1.7
Duolite C_1	2	Synthetic resin cation exchangers		0.5 to 0.55
C_2				1.115 to 1.25
Ionac A-293-M	1	Synthetic resin acid adsorbent or anion exchanger	Stable polymer with active amine groups	0.75 (breakthrough)
Ionac A-300	1	Synthetic resin acid adsorbent or anion exchanger	Stable polymer with active amine groups	H_2SO_4—1.5 (breakthrough)
Ionac C-200	1	Synthetic resin cation exchanger	Stable polymer with active SO_3H groups	0.55 (breakthrough) 0.82 total
Nalcite—MX	3	Synthetic resin cation exchanger	Modified phenol formaldehyde resin with nuclear SO_3H groups	0.52 to 0.70
Zeo-Dur	4	Processed glauconite	$K(Fe, Al)Si_2O_6$	0.125—0.150
Zeo-Karb	4	Coal derivative cation exchanger	Sulfonated coal	0.320—0.550

---

* 1 = American Cyanamid Company, New York, N. Y. 2 = Chemical Process Co., San Francisco, Calif. 3 = National Aluminate Corp., Chicago, Ill. 4 = The Permutit Co., New York, N. Y. 5 = The Resinous Products & Chemical Co., Philadelphia, Pa.

† The actual capacity will depend upon: (1) Regenerant. (2) Regenerant concentration. (3) Regenerant amount. (4) Rate of regeneration, or time of contact. (5) Direction of regeneration as compared with adsorption run. (6) Total solids in liquor treated. (7) Nature of ions present and amounts. (8) Rate of flow of liquor treated. (9) Depth of adsorbent bed.

application in the quantitative analytical techniques described at this conference by Dr. Claesson and Drs. Moore and Stein. Preliminary experiments performed by the author in the laboratory of Dr. K. G. Stern, utilizing exchange adsorbents as retardation barriers in the frontal analysis technique of Tiselius,[84] indicate that the "predictable" properties of ion exchangers may be put to good use.

The techniques of displacement analysis and even partition chromatography may be expected to yield highly interesting results with ion exchangers. As new exchange adsorbents become available through synthesis, the horizons for ion exchange techniques may be expected to broaden and further to enrich the contributions of chromatography to chemical research.

BIBLIOGRAPHY

1. **Abrahamczik, M.**
1938. Microchemie **25**: 228.
2. **Adams, B. A., & E. L. Holmes**
1935. J. Soc. Chem. Ind. **54**: 1.
3. **Albanese, A. A., & J. E. Frankston**
1945. Biol. Chem. **159**: 185.
4. **Applezweig, N.**
1944. J. Am. Chem. Soc. **66**: 1990.
5. **Applezweig, N., & M. Rice**
To be published.
6. **Applezweig, N., & S. Ronzone**
1946. Ind. Eng. Chem. **38**: 576.
7. **Archibald, R. M.**
1944. Biol. Chem. **156**: 121.
8. **Archibald, R. M.**
1945. J. Biol. Chem. **159**: 693.
9. **Austerweil, G.**
1934. Sucr. Belge **53**: 42.
10. **Austerweil, G.**
1938. British Patent 497,928 (Dec. 23).
11. **Austin, P. R.**
1943. U. S. Patent 2,316,215.
12. **Bahrdt, Z.**
1927. Anal. Chem. **70**: 109.
13. **Bhatnagar, S. S., A. N. Kapur, & M. S. Bhatnagar**
1940. J. Ind. Chem. Soc. **17**: 361.
14. **Binkley, D. W., D. W. MacCorquodale, S. A. Thayer, & E. A. Doisy**
1939. J. Biol. Chem. **130**: 219.
15. **Bird, P. G.**
1941. U. S. Patent 2,244,325 (June 3).
16. **Block, R. J.**
1942. Proc. Soc. Exp. Biol. & Med. **51**: 252.
17. **Block, R. J.**
1946. Chem. Rev. **38**: 501.
18. **Breazeale, E. L., L. T. Pierce, & J. F. Breazeale**
Unpublished data.
19. **Buc, S. R., J. H. Ford, & E. D. Wise**
1945. J. Am. Chem. Soc. **67**: 92.

20. **Buck, R. E., & H. H. Mottern**
1945. Ind. Eng. Chem. **37**: 635.
21. **Cannan, R. K.**
1944. J. Biol. Chem. **152**: 401.
22. **Cannan, R. K.**
1946. Ann. N. Y. Acad. Sci. **47**(2): 135.
23. **Cantor, S. M.**
1940. U. S. Patent 2,328,191 (March 20).
24. **Carter, H. E., *et al.***
1945. J. Biol. Chem. **160**: 337.
25. **Cerecedo, L. R., & D. J. Hennessy**
1937. J. Am. Chem. Soc. **59**: 1617.
26. **Cleaver, C. S., R. A. Hardy, & H. G. Cassidy**
1945. J. Am. Chem. Soc. **67**: 1343.
27. **Cruz Coke, E., F. Gonzales, & W. Hulsen**
1945. Science **101**: 340.
28. **Davis, L. E.**
1945. J. Phys. Chem. **49**: 473.
29. **Deitz, V. R.**
1944. Bibliography of Solid Adsorbents. Nat. Bur. Standards. Washington, D. C.
30. **Doisy, E. A., *et al.***
1945. U. S. Patent 2,373,105 (April 10).
31. **Doisy, E. A., *et al.***
1945. U. S. Patent 2,375,979 (May 15).
32. **Dubnoff, J. W.**
1941. J. Biol. Chem. **141**: 711.
33. **Ecker, E. E., & L. Pillemer**
1942. Ann. N. Y. Acad. Sci. **43**: 63.
34. **Ecker, E. E., L. Pillemer, S. Seifter, & N. Applezweig**
Unpublished data.
35. **Englis, D. T., & H. A. Fiess**
1942. Ind. Eng. Chem. **34**: 864.
36. **Englis, D. T., & H. A. Fiess**
1944. Ind. Eng. Chem. **36**: 604.
37. **Folin, O., & R. D. Bell**
1917. J. Biol. Chem. **29**.
38. **Freudenberg, K., H. Walch, & H. Molter**
1942. Naturwissenschaften **30**: 87.
39. **Griessbach, R.**
1939. Verlag Chemie. Berlin.
40. **Griessbach, R.**
1939. Angew. Chem. **52**: 215.
41. **Harm**
1896. German Patent 95,447 (June 2).
42. **Harrison, J. W. E., R. J. Myers, & D. C. Herr**
1943. J. Am. Philos. Assoc. **32**: 121.
43. **Hennessy, D. J.**
1941. Ind. Eng. Chem. Anal. Ed. **13**: 216.
44. **Hennessy, D. J., & L. R. Cerecedo**
1939. J. Am. Chem. Soc. **61**: 179.
45. **Herr, D. S.**
1945. Ind. Eng. Chem. **37**: 631.
46. **Higgins, E. B.**
1938. British Patent 489,027 (July 11).

47. Hochberg, M., D. Melnick, & B. L. Oser
1945. J. Biol. Chem. **158:** 265.
48. Holmes, Eric L.
1941. U. S. Patent 2,240,116 (April 29).
49. Jeanprost, C.
1932. Bull. Assoc. Chim. Sucr. Dist. **49:** 189.
50. Jones, M. C. K.
1942. U. S. Patent 2,280,237 (April 21).
51. Katzman, P. A., *et al.*
1943. J. Biol. Chem. **148:** 501.
52. Kemp, C. R., & F. Bandelin
1945. J. Am. Philos. Assoc. **34:** 306.
53. Lejwa, A.
1939. Biochem. Z. **256:** 236.
54. Liebig, G. F., A. P. Vanselow, & H. D. Chapman
1943. Soil Science **55:** 371.
55. Lyman, J. F., E. H. Browne, & H. E. Otting
1933. Ind. Eng. Chem. **25:** 1297.
56. Maizel, B.
1944. U. S. Patent 2,364,639 (Dec. 12).
57. Matchett, J. R., *et al.*
1944. Ind. Eng. Chem. **36:** 851.
58. McColloch, R. J., & Z. I. Kertesz
1945. J. Biol. Chem. **160:** 149.
59. McCready, R. M., & W. J. Hassid
1944. J. Am. Chem. Soc. **66:** 560.
60. Myers, P. B., & A. H. Rouse
1943. U. S. Patent 2,323,483 (July 6).
61. Myers, R. J.
1942. Advances in Colloid Science **I:** 317-351. Interscience. New York
62. Myers, R. J.
1944. U. S. Patent 2,341,329 (Feb. 8).
63. Myers, R. J., J. W. Eastes, & F. J. Myers
1941. Ind. Eng. Chem. **33:** 397.
64. Myers, R. J., J. W. Eastes, & D. Urquhart
1941. Ind. Eng. Chem. **33:** 1270.
65. Nachod, F. C., & W. Wood
1944. J. Am. Chem. Soc. **66:** 1380.
66. Oberst, M. S.
1938. J. Lab. & Clin. Med. **24:** 318.
67. Pierce, L. F., & E. L. Breazeale
1942. J. Invest. Dermatol. **252:** 249-255.
68. Platt, B. S., & G. E. Glock
1942. Biochem. J. **36:** Proc. XVIII.
69. Polis, B. D., & J. A. Reinhold
1944. J. Biol. Chem. **156.**
70. Potts, A. M., & T. F. Gallagher
1944. J. Biol. Chem. **154:** 349.
71. Rawlings, F. M., & R. W. Shafor
1942. Sugar **37:** 26.
72. Riley, R.
1940. U. S. Patent 2,226,389 (Dec. 24).
73. Ryznar, J.
1944. Ind. Eng. Chem. **36:** 821.
74. Schindler H.
1945. U. S. Patent 2,367,803.

75. **Sims, E. A. H.**
1945. J. Biol. Chem. **158:** 239.

76. **Smit, P.**
1940. U. S. Patent 2,198,393 (April 23).

77. **Steinberg, A.**
1944. Proc. Soc. Exp. Biol. & Med. **56:** 124.

78. **Sussman, S., & A. B. Mindler**
1945. Chem. Ind. **57:** 455.

79. **Sussman, S., A. B. Mindler, & W. Wood**
1945. Chem. Ind. **57:** 549.

80. **Sussman, S., F. C. Nachod, & W. Wood**
1945. Ind. Eng. Chem. **37:** 618.

81. **Thomas, H. C.**
1944. J. Am. Chem. Soc. **66:** 1664.

82. **Tiger, H. L., & J. G. Dean**
1942. U. S. Patent 2,293,954 (Aug. 25).

83. **Tiger, H. L., & S. Sussman**
1943. Ind. Eng. Chem. **35:** 186.

84. **Tiselius, A.**
1942. Advances in Colloid Science **I:** 81. Interscience. N. Y.

85. **Turba, F., M. Richter, & F. Kuchar**
1943. Naturwissenschaften **31:** 508.

86. **Walton, H. F.**
1943. J. Phys. Chem. **47:** 371.

87. **Weitz, F. W.**
1943. Sugar **38:** 26.

88. **Whitehorn, J. C.**
1923. J. Biol. Chem. **56:** 751.

89. **Whitehorn, J. C.**
1935. J. Biol. Chem. **108:** 633.

90. **Williams, K. T., & C. M. Johnson**
1944. Ind. Eng. Chem. **16:** 23.

THE SURFACE AREAS OF SOME SOLID ADSORBENTS OF POSSIBLE USE IN CHROMATOGRAPHY

By Victor R. Deitz

National Bureau of Standards, Washington, D C.

Most of the adsorbents available to research workers in chromatography are, naturally, those intended for industrial use. One problem with which all manufacturers of commercial adsorbents have to contend is the reproducibility of the product. Modern control methods are tending towards a more uniform product, but it is a fact that fluctuations do exist in the adsorptive properties of various consignments. Consequently, a research laboratory which plans to undertake a modest investigation using some one adsorbent, can best serve its future needs by purchasing about one hundred pounds of the adsorbent, mixing it well, and placing this stock in storage. Two-gallon narrow-neck rubber-stoppered glass bottles make convenient storage containers for laboratory work.

Since the commercial adsorbents are intended for industrial application, they are designed for use either in filtration through columns or in a batch process by agitation with the solution which requires subsequent filtration. This situation limits the available particle sizes. Practical aspects dictate that the particle sizes for the adsorbents used in column filtration lie in the range between 4-mesh and 40-mesh. Uniformity in particle size has not been stressed as much as it should be. In the batch process, the manufacturer would like to use as small a particle size as possible, since, thus, the adsorbent appears to best advantage in commercial application. However, the necessary subsequent filtration operation usually places a limit on the amount of very small particles that may be added. Too fine a particle size would consume too much filter aid in order to obtain the required filtration rates.

The commercial solid adsorbents of the United States of possible use in chromatography are listed in TABLE 1. The division into four groups is not claimed to be unique. Also, no claim is made for completeness. The choice of an adsorbent for use in chromatography is made primarily on an empirical basis and great ingenuity has been exercised by the various investigators in matching the adsorbent or mixture of adsorbents to a particular problem.

It has been pointed out by Drake and Ritter[1] that porous materials are characterized by two related qualities: first, a particle density which is appreciably lower than the true density of the material, and, second, a surface area which is greater than the observable geometric surface area.

TABLE 1

SOLID ADSORBENTS IN THE UNITED STATES FOR POSSIBLE USE IN CHROMATOGRAPHY*

A: Carbon Adsorbents

Name	Manufacturer
1. Activated Carbons from coal	Pittsburgh Coke and Chemical Company, Pittsburgh, Pa.
2. Blood charcoal, acid washed	Eimer and Amend, New York, N. Y.
3. Bone Char from the pyrolysis of animal bones	American Agricultural Chemical Company, Detroit, Mich. Baugh and Sons Company, Philadelphia, Pa. Consolidated Chemical Industries, New York, N. Y.
4. "Cliffchar" from wood charcoal	Cliffs-Dow Chemical Company, Marquette, Mich.
5. Columbia Activated Carbons from coconut shells and hard nuts	Carbide & Carbon Chemicals Corporation, New York, N. Y.
6. "Darco" Carbons from lignite or wood	Darco Corporation, New York, N. Y.
7. "Minchar" from carbonaceous slate	Minchar Manufacturing Co., Elmira, N. Y.
8. "Norit" Carbons from pinewood-charcoal	American Norit Company, Inc., Jacksonville, Fla.
9. "Nuchar" Carbons from paper-pulp waste liquors	West Virginia Pulp & Paper Company, New York, N. Y.
10. "Suchar" from paper-pulp waste liquors	West Virginia Pulp & Paper Company, New York, N. Y.
11. Sugar Charcoals	Laboratory Supply Firms
12. "Weschar" from lignite	Western Filter Company, Denver, Colorado

B: Mineral Adsorbents

Name	Manufacturer
1. Activated Alumina from bauxite ores	Aluminum Ore Company, East St. Louis, Ill.
2. "Alba-floc", a finely divided $CaSO_4$	U. S. Gypsum Company, Chicago, Ill.
3. Attapulgus clay products from clay deposits "Driocel", from bauxite	Attapulgus Clay Company, Philadelphia, Pa.
4. Bleaching Clay number 260	Industrial Minerals & Chemical Co., Berkeley, Calif.
5. "Celite" Filter Aids, diatomaceous earths	Johns Manville Corporation, New York, N. Y.
6. "Crystalite", a zeolite for water softening	Infilco, Inc., Chicago, Ill.
7. "Dicalite" Filter aids, diatomaceous earths	Dicalite Company, New York, N. Y.

---

* A more comprehensive tabulation may be found in: Victor R. Deitz. *Bibliography of Solid Adsorbents.* 958 pp. (J. M. Brown, Committee Chairman, Revere Sugar Refinery, 833 Medford St., Charlestown 29, Mass. 1944).

TABLE 1 (*continued*)

Name	Manufacturer
8. Filtrol Products: "Filtrol", an activated clay from bentonites of montmorillonite structure "J-Neutrol", same as "Filtrol" "Desiccite", a dehydrated bentonite	Filtrol Corporation, Los Angeles, Calif.
9. Floridin Products, bleaching earths: "Diluex", a selected "Floridin" of which 85% is finer than 300 mesh "Floraid", an adsorbent filter aid—95% 300 mesh Florida Fullers Earth, one grade called "Floridin", a natural fullers earth and another called "Florex", an extruded fullers earth* "Florigel", a hydrated Florida Fullers Earth† "Florisil", an analytical adsorbent 40-60, 60-100 mesh, also from 4 to 300 mesh "Florite", an activated bauxite in 20-60 and other meshes	Floridin Company, Inc., Warren, Pa.
10. Gypsum, hydrated	Chemical Supply Firms
11. Lloyd's reagent, prepared from fullers earth	Eli Lilly & Company, Indianapolis, Ind.
12. "Porocel", an activated bauxite	Porocel Corporation, Philadelphia, Pa.
13. "Sinclair Earth", a fullers earth	Sinclair Refining Company, New York, N. Y.
14. Talc, purified powder U.S.P.	Eimer and Amend, New York, N. Y.
15. "Volclay", a natural bentonite	American Colloid Co., Chicago, Ill.
16. "Zeo-Dur", the zeolitic green sand	Permutit Company, New York, N. Y.

C: Inorganic Compounds

1. Aluminum Oxide, anhydrous (purissimum or according to Brockman)	Merck & Company, Inc., Rahway, N. J.
2. Calcium hydroxide (lime)	Chemical Supply Firms
3. Calcium carbonate, precipitated	Chemical Supply Firms
4. "Decalso", synthetic zeolite	Permutit Company, New York, N. Y.
5. "Defluorite", a tricalcium phosphate preparation	Aluminum Corporation of America, Pittsburgh, Pa.
6. "Drierite", anhydrous $CaSO_4$	W. A. Hammond Drierite Company, Xenia, Ohio
7. "Hydralo", an activated aluminum oxide	J. T. Baker Chemical Company, Phillipsburg, N. J.
8. Magnesia	Chemical Supply Firms

* "B" 16-30 mesh; "S" 30-60 mesh; "XXS" 60-100 mesh; "XXF" 90% through 100 mesh; "XXX" 90% through 200 mesh.

† On 70 mesh 4%; 70-80, 4%; 80-100, 5%; 100-140, 14%; 140-200, 11%; 200-300, 14%; through 300, 48%.

TABLE 1 (*continued*)

Name	Manufacturer
9. (Micron Brand Magnesium oxide, #2641)	Westvaco Chlorine Products Corp., New York, N. Y.
10. Magnesium trisilicate "Magnesol"	Westvaco Chlorine Products Corp., New York, N. Y.
11. Magnesium trisilicate, "Florisil"	Floridin Company, Inc., Warren, Pa.
12. Magnesium trisilicate No. 34	Philadelphia Quartz Company, Ltd., Berkeley, Calif.
13. "Santocel", a silica aerogel	Monsanto Chemical Company, Boston, Mass.
14. "Silene EF", a synthetic, hydrated calcium acid silicate	Columbia Chemical Division, Pittsburgh Plate Glass Company, Barberton, Ohio
15. Silica gel "Protex-Sorb"	Davison Chemical Company, Baltimore, Md.

D: Organic Compounds

1. "Amberlite" ion-exchange resins: IR—1, IR—100 cation exchangers (phenol-formaldehyde) IR-4, acid-adsorbent (amine-formaldehyde)	Resinous Products & Chemical Co., Philadelphia, Pa.
2. Cellulose, many sources, one being filter paper shredded in a Waring Blendor	
3. "Deacidite", ion exchange resin	The Permutit Company, New York, N. Y.
4. "Duolites", ion exchange resins	Chemical Process Company, San Francisco, Calif.
5. Inulin	Eimer and Amend, New York, N. Y.
6. "Ionac", ion exchange resins, made from melamine compounds	American Cyanamide & Chemical Co., New York, N. Y.
7. Lactose	Eimer and Amend, New York, N. Y.
8. Sucrose—multiple X Confectioners' powdered sugar with 3% starch	Cane Sugar Refineries
9. "Zeo-Karb", a carbonaceous zeolite from coal	The Permutit Company, New York, N. Y.

This observation is true for most solid adsorbents since they, too, may be rather porous substances.

It has been amply demonstrated in a great number of recent investigations that the extent of surface is the factor of decisive importance in physical adsorption. In chemisorption, and in adsorption from solutions, the extent of surface is an important auxiliary factor in determining the progress of a chemical interaction of the surface with the adsorbate. Previous to the work of Brunauer, Emmett, and Teller,[24] the extent of surface could be determined, by the then available methods, only with an accuracy corresponding to a 2- to 4-fold order of magnitude. It is now possible, by means of the determination of adsorption of nitrogen at low temperatures, to ascertain the surface area with an absolute accuracy of

TABLE 2

SURFACE AREA AND DENSITY OF SOME SOLID ADSORBENTS

Material	Surface area		Density gms./ml.			Reference
	m.²/gm.	Gas at °C.	Bulk	Particle	True	
		A: Carbon Adsorbents				
Bone char, new	120	N_2 at-195°	0.8	—	2.88	2, 7
Coconut-shell charcoal	1700	"	—	—	2.08	2
"Darco G 60"	1300	"	—	—	—	2
"Darco S 51"	500	"	—	—	—	2
Granular "Darco"	620	"	—	—	2.14	3
"Darco" carbon	560	"	0.36	0.732	2.086	1
"Columbia" carbon	1397	"	0.42	0.754	1.897	1
Graphite, powdered	30.73	"	—	—	2.26	4, 7
Carbon blacks: (see also 26)						
"Arrow black" (.03 μ)	112	"	—	—	—	2, 4
"Acetylene black" (.05 μ)	64.5	"	—	—	—	4
"Micronex" (.03 μ)	106.7	"	—	—	—	4
"Thermatomic carbon" (.5 μ)	6.81	"	—	—	—	4
"Norit"	930	"	—	—	2.04	2, 7
"Suchar"	850	"	—	—	—	2
Petroleum coke	0.52	"	—	—	—	2
"Zeo-carb"	0.2	"	—	—	—	7
		B: Mineral and Inorganic Adsorbents				
Activated alumina, "Alorco"	200	N_2 at-195°	—	—	—	5
Activated clay -6	147	"	0.81	1.242	2.495	1
Activated clay -7	223	"	0.68	1.111	2.614	1
Activated alumina desiccant	175	"	0.81	1.547	3.675	1
Alumina, synthetic 10-30 mesh	170	"	0.88	1.58	3.23	10
Asbestine pigment	4.1	H_2O at 25°	—	—	—	6
Attapulgus clay	170	N_2 at-195°	—	—	—	7
Asbestos	18	"	—	—	—	8
Basic carbonate white lead	1.1	H_2O at 25°	—	—	—	6
Barytes pigment	0.59	"	—	—	—	6
Barium sulfate	0.43	"	—	—	—	14
Bauxite 8-14 mesh	228	N_2 at-195°	0.88	1.575	3.622	1
Bauxite, Arkansas I, 10-20 mesh	289	"	0.90	—	—	9
Bauxite, Arkansas II, 10-20 mesh	353	"	0.90	—	—	9
Bauxite, British Guiana, 10-20 mesh	280	"	0.90	—	—	9
Bauxite, French, 10-20 mesh	82.5	"	1.30	—	—	9
Bauxite, Arkansas 1, 10-30 mesh	236	N_2 at-195°	0.97	—	—	10
Bauxite, Arkansas 2, 10-30 mesh	176	"	0.93	—	—	10
Bauxite, Arkansas 3, 10-30 mesh	200	"	0.86	—	—	10

TABLE 2 (*continued*)

Material	Surface area		Density gms./ml.			Reference
	$m.^2$/gm.	Gas at °C.	Bulk	Particle	True	
Bauxite, South America, flaky 10-30 mesh	187	N_2 at-195°	0.96	—	—	10
Blanc fixe pigment	2.2	H_2O at 25°	—	—	—	6
Bentonite, (<0.3 μ)	18.7	N_2 at-195°	—	—	—	11
Brucine montmorillonite	12.3	"	—	—	—	12
Clay	10	N_2 at-195°	—	—	—	8
Copper sulfate, anhydrous	6.23	"	—	—	—	16
Chromium oxide gel	228.	"	—	—	—	16
ditto, sintered	28.3	"	—	—	—	16
Diaspore, AlO(OH)	3.9	"	—	—	—	12
Diatomaceous earth	4.2	"	0.32	0.631	2.265	1
Diatomaceous earth	<1.	"	0.29	0.470	2.327	1
Ferric oxide gel	211	C_2H_5OH	—	—	—	15
Ferric oxide gel, water aged	52	"	—	—	—	15
Fullers earth	129.	N_2 at-195°	0.55	0.860	2.660	1
Giblisite, $Al(OH)_3$	0.31	N_2 at-195°	—	—	—	12
Glass spheres, 12-micron	0.216	"	—	—	—	17
Glass, sized beads (7.2 μ)	0.55	"	—	—	2.237	14
Glass, fritted, 8-14 mesh	0.69	"	0.87	1.608	—	1
Glass, porous (Corning Glass Company)	120	"	—	—	—	13
Glaucosil	82	"	—	—	—	16
Halloysite	43.2	"	—	—	—	12
Illite (hydrous mica, < 0.3 μ)	97.1	"	—	—	—	12
Kaolinite (< 0.3 μ)	15.5	N_2 at-195°	—	—	—	12
Kieselguhr	22.2	"	—	—	—	18
Lithopone pigment	2.3	H_2O at 25°	—	—	—	6
Montmorillonite	15.5	N_2 at-195°	—	—	—	12
Montmorillonite	11.2	"	—	—	—	11
Palladium catalyst	0.35	"	—	—	—	22
Paper, insulating	0.55	"	—	—	—	28
Paris white	2.	"	—	—	—	8
Porcelain, Coors	1.6	"	1.00	1.801	2.612	1
Potassium chloride (< 200 mesh)	0.24	"	—	—	—	16
Pumice	0.38	"	—	—	—	16
Quartz, 400 mesh	0.361	"	—	—	—	17
Quartz, flint, 4-6 mesh	<1.	"	1.53	2.610	2.641	1
Silica gel	620.	"	—	—	—	3
Silica gel	765	"	—	—	—	3
Silica gel I, non-electrolyzed	584	"	—	—	—	16
Silica gel II, electrolyzed	614	"	—	—	—	16
Silica gel	866	N_2 at-195°	—	—	—	19
Silica gel	414	"	—	—	—	20
ditto + 5.2% Al_2O_3	541	"	—	—	—	20
Silica gel, 4-8 mesh	669	"	0.77	1.226	2.251	1
Silica gel, 10-30 mesh	602	"	0.78	1.32	2.12	10

TABLE 2 (*continued*)

Material	Surface area		Density gms./ml.			Reference
	m.²/gm.	Gas at °C.	Bulk	Particle	True	
Silica aerogel V	410	CCl_4	—	—	—	21
Silica aerogel V	370	C_6H_6	—	—	—	21
Silica aerogel III	330	H_2O	0.92	—	—	21
Silica aerogel	393	Butane, 0°	—	—	—	15
Silica aerogel	690	N_2 (—183°)	—	—	—	15
Silica aerogel, water-aged	186	Butane, 0°	—	—	—	15
Silica alumina gel, 4-8 mesh	280	N_2 at-195°	1.00	1.558	2.369	1
Silica alumina gel, 4-8 mesh	404	"	0.75	1.175	2.388	1
Silica alumina gel, 4-8 mesh	467	"	0.69	1.126	2.377	1
Silica alumina gel, 4-8 mesh	372	"	0.64	0.962	2.378	1
Silica alumina gel, 4-8 mesh	409	"	0.55	0.860	2.402	1
Silica alumina gel, 4-6 mesh	323	"	0.55	0.932	2.405	1
Silica alumina gel, 4-6 mesh	201	"	0.68	1.050	2.343	1
Titanium dioxide pigment	8.2	H_2O at 25°	—	—	—	6
Titanium dioxide	13.9	N_2 at-195°	—	—	—	27
Titanium dioxide	13.8	H_2O at 25°	—	—	—	27
Ultramarine blue pigment	13	"	—	—	—	6
Whiting pigment	2.6	"	—	—	—	6
Zinc oxide pigment	2.4	H_2O at 25°	—	—	—	6
Zinc oxide, pigment K-1602	3.80	N_2 at-195°	—	—	--	14
Zinc oxide, pigment F-1601	9.48	"	—	—	—	14
Zinc oxide, pigment G-1603	3.88	"	—	—	—	14
Zinc oxide, pigment KH-1604	0.658	"	—	--	—	14
Zirconium orthosilicate	2.76	"	—	—	—	14

±35 per cent and a relative accuracy of about ±5 per cent.[25] TABLE 2 contains a summary of the values for a large number of surface determinations which have been published in the last ten years. While this tabulation is not complete, it does demonstrate the keen interest initiated by the introduction of this new experimental tool.

The surface area available to nitrogen molecules may not necessarily be that available to larger molecules. Emmett and DeWitt[13] have shown, for example, that the surface areas of a porous glass as determined with nitrogen, argon, or oxygen are in agreement but that determined with butane was about 40 per cent lower. It seems reasonable to find that some parts of a porous adsorbent would be available to a small gas molecule and not to a larger molecule. Emmett,[25] however, has since indicated some justification for assigning to some adsorbate molecules larger cross-sectional areas than those calculated from liquid densities. For example, values of 11.3, 56.6, and 64 sq. Å for the water, *n*-butane, and *n*-heptane molecule, respectively, instead of 10.5, 32, and 45 sq. Å were found by Emmett to make the areas of crystalline solids calculated by Harkins and Jura agree with those calculated by the BET method. It is important to view these area determinations as a concept of surface available to the particular adsorbate molecule until more is learned about a greater variety of adsorption systems.

There are other properties besides density (bulk, particle, or true values) and surface area which are pertinent to a physical description of an adsorbent. Particle size is an important characterization. Zechmeister and Cholnoky[23] remark that the particle size of the column contents must not be too diverse, otherwise the finer particles may act almost like a different adsorbent. The particle size distribution may influence three important experimental quantities: (1) rate of flow; (2) channeling through column; and (3) sharpness of the boundaries separating zones. The dependence of the rate of flow on particle size is well known. Channeling is due to the presence of a path of least resistance to flow, and the chance of realizing such behavior is greater when the particle size distribution is broad. The reproducible packing of an adsorption column becomes a simpler operation when the particles are of uniform size. Variation in the shapes of the particle tend to cancel out on the average. It is apparent that the closer the kinetics of flow are to simple piston-displacement, the sharper may be the boundaries separating the zones. A closer control of particle size distribution is one factor to which the manufacturers of adsorbents could contribute in order to realize the standardization of materials for use in chromatography.

Relatively little progress has been made on a correlation of the chemical properties of various preparations of adsorbents. A chemical analysis of the surface layer may be only approximated from the complete chemical analyses of the adsorbent itself. In some special cases, the structure of the adsorbing surface has been worked out in detail.

In solution adsorption phenomena, it is probable that the adsorption is confined to monolayers. All components of the solution compete for the available adsorbing surface. In dilute solutions, for example, it is likely that the solvent is adsorbed in great excess to the solute. It is very desirable to know the absolute amounts of all adsorbed components, and this constitutes an outstanding problem in the present studies of adsorption from solutions.

BIBLIOGRAPHY*

1. **Drake, L. C., & H. L. Ritter**
   1945. Anal. Ed., Ind. Eng. Chem. **17**: 787-91.
2. **Deitz, V. R., & L. F. Gleysteen**
   1942. J. Research N.B.S. **29**: 191-225.
3. **Gleysteen, L. F., & V. R. Deitz**
   1945. J. Research N.B.S. **35**: 285-307.
4. **Emmett, P. H., & T. DeWitt**
   1941. Anal. Ed., Ind. Eng. Chem. **13**: 28-33.

---

* These reference numbers also appear in TABLE 2.

5. **Krieger, K. A.**
1941. J. Am. Chem. Soc. **63**: 2712-4.
6. **Smith, D., & H. Green**
1942. Anal. Ed., Ind. Eng. Chem. **14**: 382-6.
7. Unpublished work, National Bureau of Standards.
8. **Bugge, Kerlogue, & Westwick**
1946. Nature **158**: 28.
9. **LaLande, W. A., Jr., W. S. W. McCarter, & J. B. Sanborn**
1944. Ind. Eng. Chem. **36**: 99-109.
10. **Heinemann, H., K. A. Krieger, & W. S. W. McCarter**
1946. Ind. Eng. Chem. **38**: 839-42.
11. **Makower, B., T. M. Shaw, & L. T. Alexander**
1937. Proc. Soil Science Soc. America **2**: 101-8.
12. **Nelson, R. A., & S. B. Hendricks**
1943. Soil Sci. **56**: 285-96.
13. **Emmett, P. H., & T. W. DeWitt**
1943. J. Am. Chem. Soc. **65**: 1253-62.
14. **Emmett, P. H.**
1941. Proc. A.S.T.M. **41**: 95-106.
15. **Harvey, E. N.**
1946. J. Am. Chem. Soc. **65**: 2343-6.
16. **Brunauer, S., & P. H. Emmett**
1937. J. Am. Chem. Soc. **59**: 2682-9.
17. **Gaudin, A. M., & F. W. Bowdish**
1944. Mining Technology A.I.M.M.E. **8**: No. 1666.
18. **Anderson, R. B.**
1946. J. Am. Chem. Soc. **68**: 688-91.
19. **Joyner, L. G., E. B. Weinberger, & C. W. Montgomery**
1945. J. Am. Chem. Soc. **67**: 2182-8.
20. **Elkin, P. B., C. G. Shull, & L. C. Roess**
1945. Ind. Eng. Chem. **37**: 327-30.
21. **Kistler, S. S., E. A. Fischer, & I. R. Freeman**
1943. J. Am. Chem. Soc. **65**: 1909-19.
22. **Emmett, P. H., & N. Skau**
1943. J. Am. Chem. Soc. **65**: 1029-35.
23. **Zechmeister, L., & L. Cholnoky**
1941. *Principles and Practice of Chromatography.* English translation by A. L. Bacharach and F. A. Robinson. Chapman & Hall, Ltd. London.
24. **Brunauer, S., P. H. Emmett, & E. Teller**
1938. J. Am. Chem. Soc. **60**: 309-19.
25. **Emmett, P. H.**
1946. J. Am. Chem. Soc. **68**: 1784-9.
26. **Smith, W. R., F. S. Thornhill, & R. I. Bray**
1941. Ind. Eng. Chem. **33**: 1303-7.
27. **Jura, G., & W. D. Harkins**
1944. J. Am. Chem. Soc. **66**: 1356-62.
28. **Stamm, A. J., & M. A. Millett**
1941. J. Phys. Chem. **45**: 43-54.

CONCLUDING REMARKS

By Harold G. Cassidy

Sterling Chemistry Laboratory, Yale University, New Haven, Connecticut

The foregoing papers are a partial record of the Conference on Chromatography held by The New York Academy of Sciences. The record is a partial one for two reasons. Firstly, it has been impossible to include an interesting paper on *Chromatography in the Isolation and Separation of the Various Penicillins* by Dr. Jacques L. Wachtel, though Dr. Wachtel was able to speak freely on this subject and to hold some discussion of it. Secondly, the record is a partial one because no written words can do justice to the discussion, to the stimulation of new ideas, or to the clarifications of old problems which come from such a meeting.

In gathering the material for this monograph, the Chairman, with the full support of the Editors, has sought partially to redress the second source of incompleteness mentioned above. It was apparent from the discussion of papers that at least three additional topics needed attention. It seemed appropriate to invite Dr. Victor R. Deitz to write a paper in which data characterizing solid adsorbents of possible use in chromatography were gathered into handy, usable form. This paper should facilitate a quantitative approach to certain of the problems of chromatography.

Ion-exchange had not been brought formally under the scrutiny of the Conference, yet it is certainly assuming increasing importance as a tool which may be applied chromatographically. Accordingly, at the suggestion of one of the Editors, Mr. Norman Applezweig was invited to write the paper reviewing this field. It is hoped that this paper will aid in bringing before a wider audience the natural and synthetic exchange substances.

The non-uniformity of adsorbent surfaces was a subject of uniform concern and frequent discussion at the Conference. It therefore seemed proper to the Chairman to invite Dr. Leo Shedlovsky to contribute a paper on separations using foams and emulsions. The surfaces involved in mobile interfaces are likely to be of a sameness throughout the system, and, thus, to meet some of the requirements hoped for in the discussion and be worthy of further study. Moreover, while chromatography has concerned itself almost entirely with solid-liquid or solid-gas interfaces, it has seemed to the Chairman that liquid-liquid or liquid-gas interfaces should also come into consideration here. The experiments of Schütz and others have now shown that interesting separations can be obtained by the counter-current application of partition between the liquid-gas interfacial and bulk liquid phases. This paper by Dr. Shedlovsky may stimulate chromatographers to further investigation of the mobile interfaces.

It is perhaps worth while to report the general pattern which seemed, to the Chairman, to run through the Conference.

Chromatography, which was first devised as a tool applied empirically, has in the last decade or so become better understood. With understanding has come the flourishing development described by Dr. Zechmeister in his first paper. We are in the process, now, of further understanding and sharpening not only the classical Tswett chromatography but also the elegant simplifications introduced by Tiselius and Claesson, and the ingenious and elegant "partition chromatography" of Martin and Synge.

It is probably correct to say that, while it may look as though some ultimate point has been reached in the refinement and subtlety of chromatography as these have been revealed in the foregoing papers, yet experience allows us to prophesy that there will be no end to progress. It is quite reasonable to expect that some of those who attended the Conference or will read these papers will be stimulated to important new ideas.

One area of investigation in which there is evident need for continued research and inspiration—continued, because research and inspiration are immanent in the foregoing papers—is that which connects adsorbability with chemical or physical properties of molecules. It was implied, in the discussions, that we should expect no simple relationship to be found here. Yet I venture to wonder if we might not, through some principle of exclusion, find a simple set of relations. We may find that, under given circumstances, a large proportion of the factors which might theoretically add complexity to this relationship drop out instead, leaving only one or two controlling factors to be dealt with.

A basis for such a hope as this may be found in the work of Tiselius and his group. Here, in the face of the complexities of chromatographic theory so well delineated by Dr. Thomas, simplifications were wrought which have brought the theory under practical control. An example of such simplification is the use of the displacement developer, which with one stroke can control the behavior of the rest of the adsorbates in the column. Perhaps, with this precedent, we can look for further simplifications, whether in technique or theory or in relations which connect structure and properties of molecules with their behavior in the column.

The need for standardized adsorbents, which was stressed especially by Dr. Zechmeister, has already been mentioned.

Those who attended the Conference and who will read these pages are indebted to the authors. They gave freely of their time, and, in some cases, came from far away to contribute to the success of the Conference and the existence of this monograph. Theirs was a real act of faith in the objective of these conferences: The free interchange of scientific knowledge and ideas among scientists. The spirit which animated the Conference was in accord with the advice given by Bacon long ago: "We advise all men to think of the true ends of knowledge, and that they endeavor not after it for curiosity, contention, or the sake of despising others, nor yet for reputation or power or any other such inferior consideration, but solely for the occasion and uses of life."

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

VOLUME XLIX. ART. 3. PAGES 327-482

APRIL 15. 1948

BIOLUMINESCENCE*

By

E. NEWTON HARVEY, RUBERT S. ANDERSON, JOHN B. BUCK, AURIN M. CHASE, HENRY EYRING, AND FRANK H. JOHNSON

CONTENTS

Introductory Remarks: A General Survey of Bioluminescence. *By* E. NEWTON HARVEY 329

Chemiluminescence in Aqueous Solutions. *By* RUBERT S. ANDERSON 337

The Chemistry of *Cypridina* Luciferin. *By* AURIN M. CHASE 353

The Fundamental Action of Pressure, Temperature, and Drugs on Enzymes, as Revealed by Bacterial Luminescence. *By* FRANK H. JOHNSON AND HENRY EYRING 376

The Anatomy and Physiology of the Light Organ in Fireflies. *By* JOHN B. BUCK 397

---

* This series of papers is the result of a Conference on Bioluminescence, held by the Section of Biology of The New York Academy of Sciences, November 8, 1946.

Publication made possible through grants from the John Strong Newberry Fund, the Esther Herrman Fund, and the General Funds of The New York Academy of Sciences.

COPYRIGHT 1948

BY

THE NEW YORK ACADEMY OF SCIENCES

INTRODUCTORY REMARKS: A GENERAL SURVEY OF BIOLUMINESCENCE

By E. Newton Harvey

Osborn Professor of Biology, Princeton University, Princeton, New Jersey

Bioluminescence occupies a somewhat paradoxical position among biological subjects. It is not only a specialized and restricted field of inquiry but at the same time a very broad one. It is restricted in that relatively few of the very large number of animal or plant groups have developed the singular ability to produce light, and broad in that it presents problems of the greatest interest to every branch of biology—to the systematist, histologist, morphologist, physiologist, biochemist, and ecologist as well as to the student of animal behavior and of evolution. Furthermore, the beauty and the mystery of animal light have always aroused the interest of the traveler and the student of other fields such as pure chemistry and physics.

Among the great chemists and physicists of the past who have taken a more than casual interest in bioluminescence are Boyle, Newton, Franklin, Priestly, Réaumur, Dessaignes, Becquerel, Davy, and Faraday. Many naturalists have been fascinated by luminous organisms, and it is rather surprising to find that Darwin, despite the variety of his travels and the breadth of his interest, casually mentions luminous animals and draws no inferences for the theory of natural selection from the widespread ability of many living forms to emit a cold light. Among biologists who have made a special study of animal light may be mentioned the names of Anderson, Beijerinck, Buck, Chase, Dahlgren, Dubois, Ehrenberg, Eyring, Giese, Harvey, Heinrich, Heller, Johnson, Kanda, Kishitani, Kluyver, Krukenberg, Mangold, McDermott, McElroy, Okada, Panceri, Phipson, Pierantoni, Pratje, Quatrefages, Spallanzani, Tilesius, Trojan, van der Burg, van der Kirk, van Schouwenburg, and Zirpolo. We regret that more of our contemporaries in this group were not present at the conference.

About 40 different orders of animals are self-luminous. These include such diverse organisms as bacteria, at least two groups of fungi, radiolaria, dinoflagellates and cystoflagellates; sponges, hydroids, medusae, siphonophores, pennatulids, ctenophores, and nemerteans; 7 families of marine worms, earthworms; ostracod, copepod, decapod and schizopod crustaceans: myriapods, possibly spiders; spring-tails, flies, and beetles; brittle-stars, bivalves, nudibranchs, and two orders of squid; balanoglossids, ascidians, and several orders of fish, both elasmobranchs and teleosts. The phyla Platyhelminthes, Nemathelminthes, Trochelminthes and all vertebrates above the fish (amphibians, reptiles, birds, and mammals) are non-luminous. All higher plants (Bryophytes, Pteridophytes, and Spermatophytes) also contain no luminous forms.

Among the luminous groups listed above, only terrestrial or marine animals produce light, while fresh-water organisms do not, even though closely related to luminous marine species. The nearest approach to a luminous fresh-water animal is an aquatic firefly larva breathing by

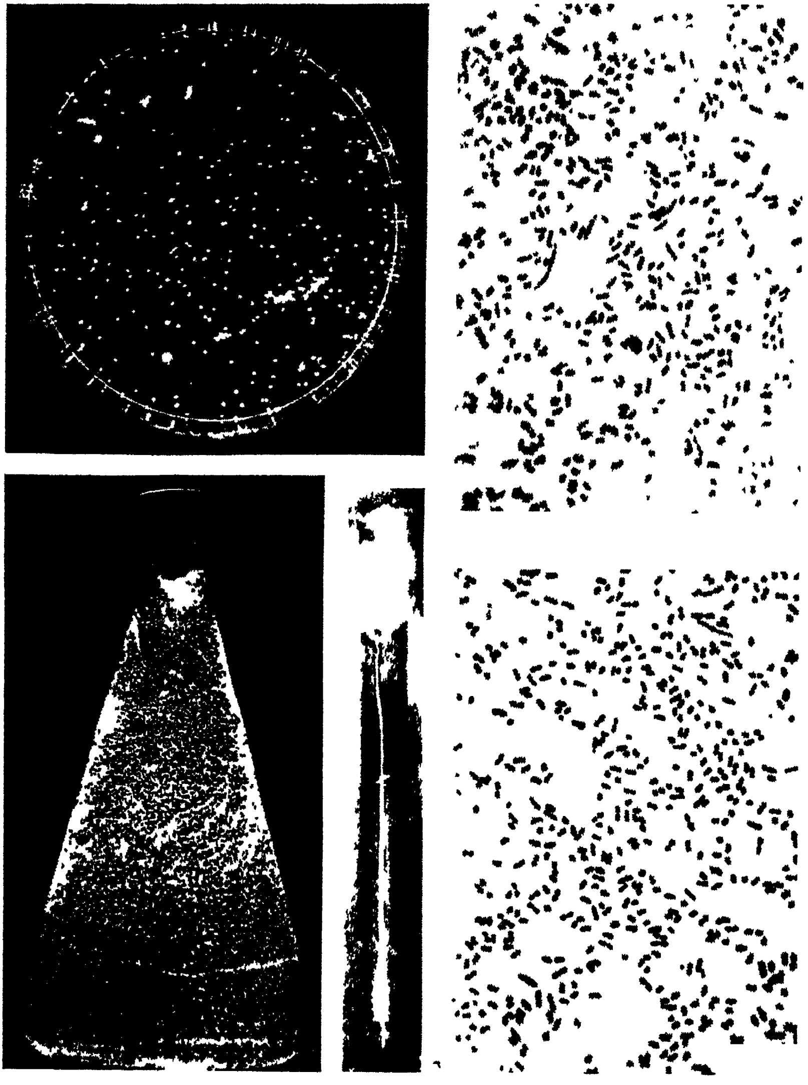


FIGURE 1. Left, luminous bacterial colonies growing on a Petri dish, flask, and test tube of culture medium (after Molish). Right, luminous bacteria highly magnified, *Pseudomonas toyamensis* (above) and *Coccobacillus ikiensis* (below), after *Kishitani*.

tracheal gills. The diverse and apparently chance distribution of luminescence in the living world suggests that the mechanism for producing light has developed in the course of evolution from some one of the chemical systems already generally present in cells and probably from one concerned with cell respiration.

Hence, much is to be gained from a study of luminescence in organisms such as luminous bacteria, where the cell respiration is intimately connected with the production of light. These bacteria occur in the sea and frequently form colonies on dead fish or squid and also on meat in refrigerators. They are of several species, are non-pathogenic to humans, and easy to culture on 2 per cent peptone, 1 per cent glycerine sea water agar at room temperature or below. A pinch of $CaCO_3$ can be added to maintain the proper pH. A good idea of their appearance can be obtained from FIGURE 1.*

In addition to such saprophytic bacteria, there are also parasitic forms. Various organisms, such as sand fleas, shrimps, midges, or caterpillars may become spontaneously infected with parasitic luminous bacteria and develop a luminous malady that is finally fatal. In the meantime, the host animals move about and would be mistaken for true luminous organisms if the origin of the light were not known.

Finally, certain squid and fish may be occasionally or regularly luminous because of the harboring of luminous bacteria in their glands. No harm results to the host from these bacteria. In the remarkable cases of the East Indian fish, *Photoblepharon* and *Anomalops* (shown in FIGURE 2), a special light organ has been developed under the eye in which luminous bacteria are always present. They live between special long cylindrical cells which are richly supplied with blood capillaries. Moreover, the continuous light emission of the bacteria can be shut off by a screening mechanism of the fish. In *Photoblepharon,* there is a fold of black membrane, like an eyelid, that can be drawn up over the organ, thus obscuring the light. In *Anomalops,* the light organ is attached at the anterodorsal corner by a hinge that allows the whole organ to be turned over and downward into a black-pigmented groove or pocket, so that none of the light-emitting surface is visible.

These two fish alone present enough problems for the evolutionists. Why should these two very closely related genera have developed two totally different methods of shutting off the light? When do the bacteria get into the organ? Nothing is known about the embryology of the fish or why so large a light surface should appear immediately under the eye. Such problems might warrant an expedition to the Banda Sea, the only

---

* FIGURES 1-4 are reproduced from *Living Light* by E. Newton Harvey, courtesy of Princeton University Press.

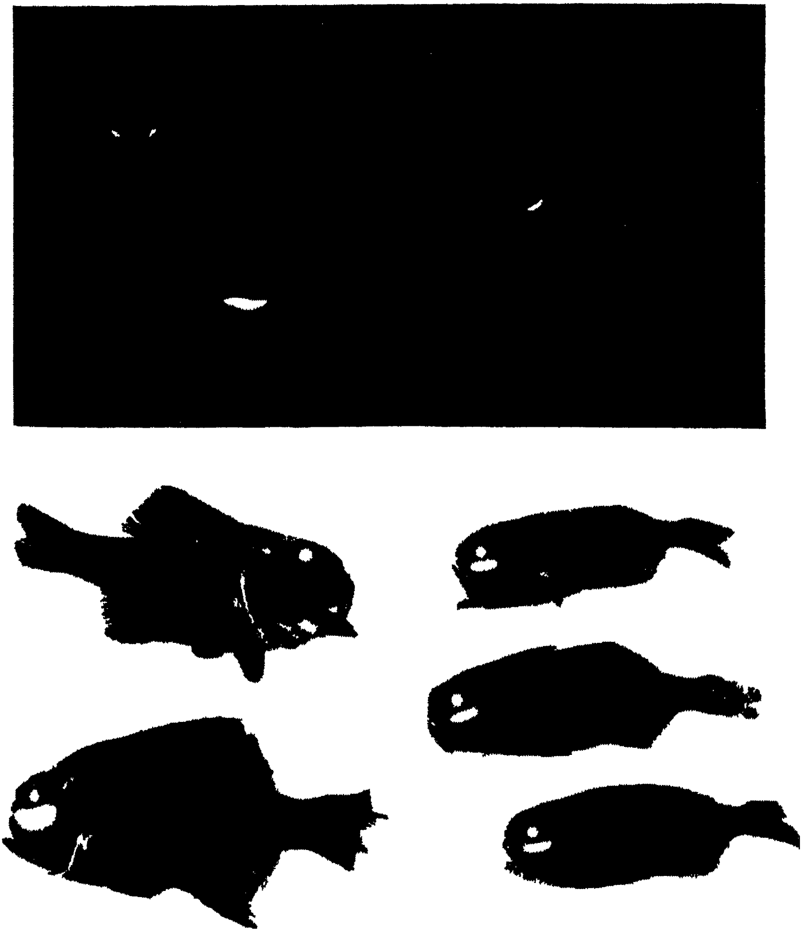
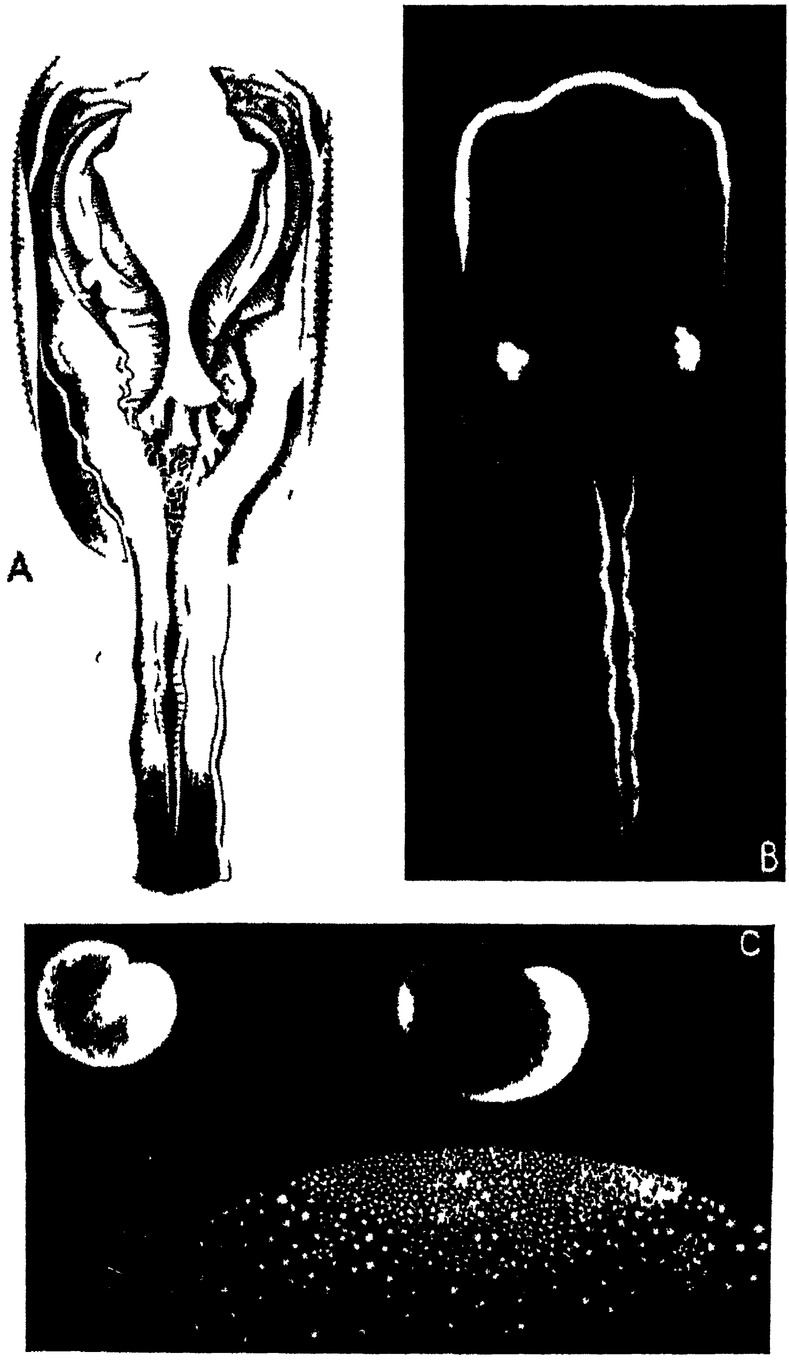


FIGURE 2 Above the fish *Photoblepharon palpebratus* swimming in water of the Banda Islands at night (After Dahlgren from a painting by Bruce Horsfall) Below *Photoblepharon palpebratus* (left) and *Anomalops katoptron* (right) photographed from dead specimens showing the large luminous organ under each eye whose light comes from symbiotic luminous bacteria

---

(See opposite page)

FIGURE 3 A *Pholas dactylus* the luminous clam with which Raphael Dubois demonstrated the luciferin luciferase reaction in daylight B The same at night to show the luminous regions (after Panceri) C The flagellate *Noctiluca miliaris* showing luminescence of two whole organisms and the appearance under the microscope where the light can be seen to come from small discrete luminous granules (after Quatrefages)


A
B
C

place in the world where *Photoblepharon* and *Anomalops* occur in considerable numbers.

Bioluminescence is actually a chemiluminescence, *i.e.*, light production during a chemical reaction. It is quite fitting, therefore, that a conference on bioluminescence should begin with light production by pure organic substances in solution. Many such compounds are known and the light emitted by some is indeed brilliant. Examples of these chemiluminescences will be discussed by Dr. Anderson.

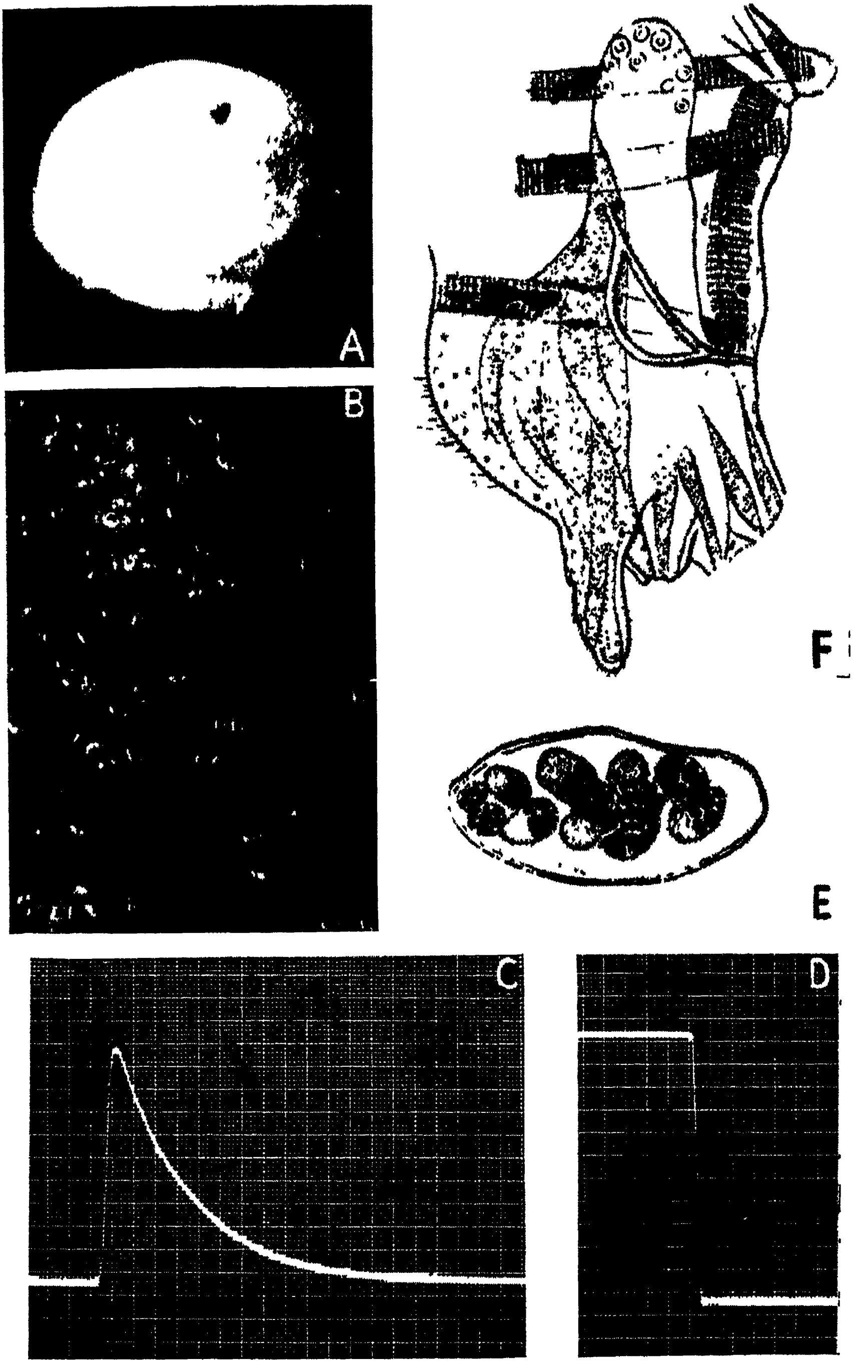
The chemiluminescences responsible for the light of organisms involve the oxidation of a compound, luciferin, in presence of an enzyme, luciferase. In some organisms (fireflies and bacteria) the oxidation is intracellular; in others (*Cypridina* and *Pholas*) extracellular, a large amount of luminous secretion being stored in a special gland. In certain organisms (bacteria, protozoa), a single cell may produce the light. In others (shrimp, squid, fish), accessory structures have been developed, so that the light organ can be truly described as a lantern, with lens, reflector, and sometimes both color and opaque screen. In the bacteria and fungi, the light emission is continuous day and night; whereas in all other forms. it appears only on stimulation.

Success in the study of any biological phenomenon is dependent on particularly favorable experimental material, *i.e.*, what we have come to speak of as classic forms. A good example of classic organisms for study are the luminous bacteria which have already been mentioned. Their contribution to a study of bioluminescence will be discussed by Dr. Johnson. Unfortunately, no one has as yet succeeded in extracting luciferin and luciferase from luminous bacteria. These substances, first demonstrated by Raphael Dubois in the elaterid beetle, *Pyrophorus,* in 1886. and later obtained from the mollusc, *Pholas dactylus,* and studied in detail by Dubois, are best extracted from luminous animals with extracellular luminescence, where a large quantity of luminous secretion is formed. *Pholas* is pictured in FIGURE 3A and B.

In 1916, during a trip to Japan for the collection of luminous squid. I noticed the abundant luminous secretion of the small ostracod crustacean, *Cypridina hilgendorfii,* and was at once able to establish the presence of luciferin and luciferase. These substances are secreted from separate. long, single, gland cells opening at pores near the mouth. Moreover.

(*See opposite page*)

FIGURE 4. A. A single ostracod crustacean, *Cypridina hilgendorfii,* enlarged, showing the black eye-spot, tip of swimming legs (below) and protuberances of luminous gland (right). B. Dried *Cypridinae,* life-size, photographed on a cloth. C. A photo-cell and string galvanometer record of light, intensely (vertical), as a function of time (horizontal), when *Cypridina* luciferin and luciferase are mixed. D. A similar record of ink mixed with water to determine time of mixing. Light intensity in arbitrary units. Each large division represents 0.2 seconds. (After Harvey and Snell). E. A cross-section of the gland of *Cypridina.* F. A longitudinal section of the gland region, showing two types of single gland cells, opening by separate pores, and cross-striated muscle fibers, whose contraction squeezes the secretion into the sea water (after Yatsu).


A
B
F
E
C
D

the whole animal, less than an eighth of an inch long, can be obtained in fair quantities and, when dried rapidly, indefinitely retains the ability to luminesce whenever moistened. Such dried material has been invaluable in a study of the chemistry of luciferin, which, together with some observations on *Cypridina* luciferase, will be presented by Dr. Chase. *Cypridina,* shown in FIGURE 4A and B, has become the classic organism for chemical study of bioluminescence. The histology of the gland is pictured in FIGURE 4E and F, and a record of *Cypridina* luminescence in FIGURE 4C and D.

One of the almost universal characteristics of luminous forms (except for the bacteria and fungi) is the ability to flash on stimulation. The effect is particularly well seen in the "phosphorescence of the sea", where waves dashing on shore, the splash of oars, or the wake of a boat are outlined by the flash of thousands of small marine organisms, mostly flagellates, stimulated mechanically by the motion of the water. *Noctiluca,* the best known of the flagellates, is shown in FIGURE 3C.

Higher organisms, also, possess a mechanism for turning their light on and off, well seen in the firefly flash. In fact, fireflies are classic material, not only for study of the physiological problem of flashing, but also for investigation of the fine histological details of luminous organ structure, which is necessary for an understanding of the physiology of flashing. Dr. Buck will consider these aspects of the bioluminescence problem and will also, I hope, describe his ecological studies on the firefly and the use of the flash as a mating signal. We need a scientific explanation of the use of the light in many luminous animals as well as of the mechanism of cold light emission itself, that unique phenomenon which, in the form of the fluorescent lamp, is now used to light our homes.

CHEMILUMINESCENCE IN AQUEOUS SOLUTIONS

By RUBERT S. ANDERSON

Department of Physiology, Schools of Medicine and Dentistry,
*University of Maryland, Baltimore, Maryland**

In trying to analyze a biological process such as bioluminescence, a beginning is often made by relating the biological observations to the most similar parts of pure chemistry and physics. This has been done in connection with light given off by living organisms or by extracts from them, and at least the most common type of luminescence has been generally agreed upon. In the case of the best known system, *Cypridina* luciferin and luciferase, the light is clearly the result of a chemiluminescence. Even the reaction producing it, an oxidation by oxygen, is a common type of reaction producing chemiluminescence. The bioluminescence of many forms seems to be similar, in various ways, to that of *Cypridina,* and these forms are therefore believed to contain chemiluminescent systems even though they may never have been separated from the cell.

It would be helpful if there existed a well developed body of information about the chemiluminescence of known compounds in solution. Unfortunately this is not the case, and answers to most of the questions which might arise in the study of bioluminescence cannot be given. It is the purpose of the present review to collect a number of the many scattered observations made, particularly in recent years, on chemiluminescence in water solutions, with little reference to the biological systems.

From early times, the characteristic of bioluminescence and chemiluminescence which has seemed most striking has been light-emission at ordinary temperatures, in contrast to its usual association with high temperatures. This property is also shown by fluorescence. The actual emission process is probably analogous. It results, most immediately, from loss of a quantum of energy during the change of an excited electronic state to the ground state or some other state having less energy than the initial excited one. The major difference between chemiluminescence and fluorescence is, then, the means by which the excited electronic state is first produced. In a chemiluminescent reaction, a part of the energy resulting from the reaction is channeled in a very particular way instead of being dissipated as heat. In fluorescence also, the energy is introduced into the molecule in a specific way, *i.e.,* by the absorption of radiation. Neither the emission of the radiation, nor the production of the excited state, need depend on a high temperature.

Even though excited states have been produced, in the one case by a chemical reaction and in the other by absorption of radiation, the excess energy need not be lost as radiation. While some of the molecules having the excess energy may emit light, others may lose it as heat or in chemical reaction. The ratio of the number of light emitters to those experiencing energy loss in other ways can be modified by various means, particularly by the addition of so-called quenching agents, which decrease the amount of light. Quantitatively, light emission at ordinary temperatures is a relatively uncommon process. For instance, absorption of light, capable of raising electrons to excited states, occurs at least in all colored solutions. Yet many colored solutions are not fluorescent at all and, of those that are, the efficiency rarely approaches one.

---

* Present address: Department of Physiology and Pharmacology, University of South Dakota, Vermillion, S. Dak.

The original production of the excited state from a chemical reaction is supposed to result from a crossing or near approach of potential energy surfaces. Eyring *et al.*[1] also consider that it may occur as a result of direct excitation by black body radiation, at least in the case of some extremely faint instances of chemiluminescence studied by Audubert.[2] The crossing or near approach of potential energy surfaces may perhaps be visualized in some way such as the following. If, for a compound, some configuration of the nuclei exists, a distance in the diatomic case in which the system has approximately the same potential energy with its electrons in the ground state and in an excited state, a shift may occur from the ground state to the excited state. The system may be brought to this critical configuration by a chemical reaction. When the configuration changes, the electron excitation energy is trapped and may then be lost as a quantum of light.

Considering, now, specific systems, it was stated above that no systems in solution are well understood. In fact, until fifteen or twenty years ago no reactions of specific compounds in aqueous or other solvents were known which gave off as much light for comparable amounts of material as the extracts from organisms. Lophine and the Grignard compounds probably approached them most closely. That is no longer true, since the description of the bright chemiluminescence of aminophthalhydrazide by Albrecht in 1928[3] and of dimethylbiacridinium nitrate by Gleu and Petsch in 1935.[4]

To find a chemiluminescent reaction which can be described with some completeness, it is necessary to turn to a much simpler system, a gas reaction. Gas reactions seem remote from bioluminescence or even chemiluminescence in solution, but they may act as a guide in a number of ways, and some of the same general experiments apply. In 1928 Polanyi and his co-workers reported extensive experimental work on the reaction of the alkali metals with several halogens and halides. For our purpose, it is necessary to consider only one of these reactions, that between sodium vapor and chlorine.[5] When these two elements react under the proper conditions, light is emitted. The sodium vapor is introduced at one end of a tube at a pressure of only 0.001 or 0.01 mm. Hg. The chlorine is introduced into the tube as a jet. The simplest condition is to have the sodium in excess. The reaction occurs where the two gases meet. From the character of the emitted light, the effect of pressure of reactants, the effect of foreign gases, and from the position of the luminescence or flame with respect to the deposit of reaction product (NaCl), the several steps of the reaction have been obtained. The following major events are believed to occur:

(1) $Na + Cl_2 \rightarrow NaCl + Cl$.
(2) $Na_2 + Cl \rightarrow NaCl^* + Na$.
(3) $NaCl^* + Na \rightarrow NaCl + Na^*$.
$Na^* \rightarrow Na + \text{light}$.
(4) $Na + Cl \rightarrow NaCl$.
(5) $NaCl^* \rightarrow NaCl + \text{heat}$.

Most of the light which is emitted has been identified as the D-line of sodium. This immediately shows the emitting material to be primarily sodium atoms. The frequency of the D-line is equivalent to 48.5 Cals. Reaction 1 furnishes only 35 Cals., which is not enough to produce the light emitted. Reaction 2, giving 75 Cals., provides more than the necessary energy for the emitted light and is believed to be its source. That the sodium atom produced in reaction 2 is not excited as it is produced, follows from a number of additional facts. The chemiluminescence of the reaction is quenched by foreign gases more strongly than the resonance emission of the D-line of sodium. This means that the excess energy is retained long enough for collisions to occur and dissipate the energy in some other way. The emission of the D-line by excited sodium occurs after about 10^{-8} seconds. It is, therefore, suggested by Polanyi that the excess energy is retained by some other material for a longer period than 10^{-8} seconds. This material is believed to be the sodium chloride. The excess energy of the sodium chloride molecules is then supposedly transferred to sodium atoms during collisions as indicated in reaction 3. The excited sodium atom thereupon loses the energy as light. Reactions 4 and 5 are competing non-luminescent reactions occurring at appreciable rates only at the wall. The very large effect of changes in the pressure of sodium vapor results from differential effects on the several reactions; that is, increased pressure of sodium favors reactions 2 and 3 as compared to 4 and 5, and an increased yield of light results. An increase of sodium from 0.001 to 0.01 mm. Hg. pressure increases the quantum yield, per two atoms of sodium reacting, from about 1 per cent to about 35 per cent. The calculated maximum quantum efficiency with still greater pressures of sodium was set at about 0.85, although this was not realized experimentally.

A fairly complete description, therefore, can be given of the chemical reactions leading up to the emission of the light and the side reactions which compete with those responsible for chemiluminescence. Even in this apparently very simple case, quite a number of complicating factors are present. In view of these complications in a simple gas reaction, it is not surprising that the description of events in the reaction of a large organic molecule in solution is so inadequate. However, although some of the features of these gas reactions are special to them, the general experimental approach through spectral emission curves, concentrations of reactants in relation to light emitted, quenching, etc., are applicable to studies in solution.

The number of known chemiluminescent reactions in solution is large, although this number is a very small fraction of the total known reactions. Harvey, in his book[6] indicates many of these reactions. Oxidation is a usual type of reaction, and pyrogallol, lophine, and the Grignard compounds are some of the better known materials undergoing reaction. A number of common systems previously considered to be non-radiating have been studied by Audubert.[2] By the use of specially designed counters, he reports finding very low intensities of ultraviolet light associated with many reactions.

This discussion will omit those reactions which show a very low efficiency and consider the two which seem most comparable, in amount of light emitted, to the gas reaction and to the bioluminescences. This distinction is, of course, arbitrary to some extent, since an intense chemiluminescence or fluorescence may be extinguished by a change of conditions. Recognizing this, it still seems that for some experimental purposes more emphasis should be placed on the quantitative differences between various reactions under the most favorable conditions known. In most of the known systems, the actual fraction of the total number of reacting molecules which give off light is extremely small. The fact that they have been observed at all is due in many examples to the extreme sensitivity of the eye as a detecting instrument. It may not be a case even of one in thousands or tens of thousands, but one in millions of molecules. In the reactions studied by Audubert, only one in 10^{14} or 10^{15} of the molecules which reacted emitted radiation. The aminophthalhydrazide[3] and the dimethylbiacridinium compounds[4] mentioned earlier are, in aqueous solutions, the most efficient light-producing systems which have so far been discovered. Each will be considered in some detail. Although a considerable study has been made of these reactions, definitive results upon which all authors are in agreement have not appeared.

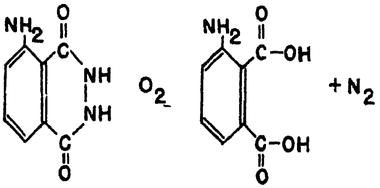
The Phthalcyclohydrazides

Albrecht,[3] in 1928, studied extensively the chemiluminescence of several phthalcyclohydrazides of which the brightest and most studied was 3-aminophthalcyclohydrazide.* This hydrazide emits a blue light when it is oxidized in alkaline solution with a number of oxidizing agents such as hydrogen peroxide, potassium ferricyanide, or sodium hypochlorite. Convenient methods of preparing the hydrazide for demonstrating the reaction[7] and additional studies of related compounds[8] appeared a few years after Albrecht's work.

---

* Also called luminol and, by Chemical Abstracts, 5-amino-2,3-dihydro-1,4-phthalazinedione.

Albrecht formulated the overall reaction as:



The time course of the reaction is markedly dependent upon the conditions and especially upon the oxidizing agent which is used. Albrecht[3] found that, with most oxidizing agents, the light is dim and of short duration. With hydrogen peroxide, the reaction lasts for a much longer time although the light intensity is low. The intensity of the light can be much increased, although the duration is decreased, by having present in addition to the hydrogen peroxide, ferricyanide, hypochlorite, or one of a number of compounds which appear to act catalytically. In general, hydrogen peroxide in the medium is necessary for the large yields of light. Other authors[6, 10–20] have subsequently studied a variety of catalysts. Manganese dioxide, colloidal platinum, hemin or hemoglobin, many metal complexes, and ozone are a few of the materials which produce increased light intensity. Weber and his collaborators more recently have studied the relation of other trivalent iron complexes,[21] copper complexes,[22] RuCl$_3$ and VOSO$_4$[23] to the chemiluminescent reaction. Whether the compound is to be looked upon as a reactant or a catalyst depends, according to these authors, upon its rate of reoxidation by atmospheric oxygen.[21]

The form of the relation between intensity of light and time depends on the catalyst used. For instance, Weber *et al.*[21] consider that the light intensity decreases exponentially with time with a number of iron catalysts. Stross and Branch,[24] using hydrogen peroxide and ferricyanide, find the curve of log intensity against time to be convex upward. They also observe that the initial point at 0.2 seconds lies below the smooth curve and suggest that the luminescent reaction may not be the initial one. The intensity-time relation of the uncatalyzed, and hence slow, reaction of hydrogen peroxide with the hydrazide depends on the concentration of hydrogen peroxide and alkali.[25, 26] Many such curves show a point of inflection because the intensity of luminescence does not decrease as rapidly during the initial portion of the reaction as it does later

It has long been known[3, 10] that the hydrazide would show visible luminescence at extremely great dilutions, one part in 10^8 parts of water. The quantitative efficiency of the light emission in terms of amount of light per mole of hydrazide present has also been studied. Harris and Parker,[27] by means of a flowing system, studied the light emitted by the hydrazide in a medium of 0.35 M sodium hydroxide, hydrogen peroxide, and sodium hypochlorite. The reaction was so fast that it was confined to a spot of light even in the flowing system. They found the efficiency to be markedly sensitive to the concentration of hydrogen peroxide, with the best condition approximately 4 moles of hydrogen peroxide per mole of hydrazide. Under their optimum conditions, from 0.003 to 0.005 as many quanta of light were produced as there were hydrazide molecules present, assuming that all of the light was emitted at 4250 Å. Use of potassium ferricyanide instead of the hypochlorite gave a more diffuse spot of light, presumably because the reaction rate was slower.

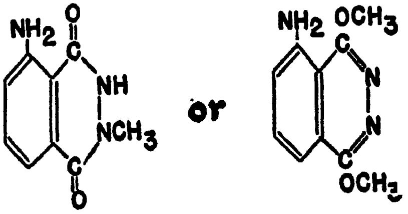
Stross and Branch[24] later published a study of the hydrazide also in a flowing system. They used varying concentrations of sodium hydroxide, hydrogen peroxide, and potassium ferricyanide. Instead of studying the light emission from a spot, they measured the light intensity at varying distances from the point of mixing. The maximum number of quanta observed per molecule of hydrazide, assuming that the light was emitted at 4570 Å, was 0.36. This is 50—100 times as great as the value given by Harris and Parker. Stross and Branch do not refer to the earlier work, so that no explanation of the discrepancy is available, apart from the different oxidizing agents.

The emitted light covers a broad region of wavelengths with a suggestion of a double maximum.[3, 28] The precise location of the curve is variously reported and may depend on the oxidizing agent or catalyst used with hydrogen peroxide. For instance, Eymers and Van Schouwenberg[28] report the maximum at 4400—4600 Å with hemin, while with ferricyanide the emission extends farther into the blue.

The emitted light, unlike the gas reaction of sodium and chlorine, gives no immediate clue as to the identity of the emitting molecule. An attempt has been made, however, to obtain some information by comparing the chemiluminescence and the fluorescence of the hydrazide and other compounds. When the fluorescence emission curve of a compound is similar to the chemiluminescence emission curve, it has been considered as evidence supporting the idea that this same compound also emits the light in the chemiluminescent reaction. Although it is not certain that the same molecule producing the two types of luminescence would necessarily emit identical spectral distribution curves, it has been considered a plausible assumption with which to begin. Such a comparison was first made by Albrecht.[3] An immediate complicating factor appears for the hydrazide. The chemiluminescence occurs best in alkaline solution, while the fluorescence appears in neutral or acid solution. Albrecht considered the two spectral distribution curves to be similar, even though the maxima were about 200 Å apart. This difference he ascribed to the different pH's of the solutions. His view was supported, in part, by Sveshnikov,[25] who brought the maxima closer together by bringing the pH's of the chemiluminescent and fluorescent solutions together more closely. The comparison of fluorescence and chemiluminescence of the hydrazide has been studied also by Eymers and Van Schouwenberg,[28] who believe that the two curves are essentially different. A comparison has also been made by Albrecht of the chemiluminescence and fluorescence of other phthalcyclohydrazides. He stated that the color of the emitted light shifted in a parallel manner in a series of derivatives. Whatever view may be the correct one, most discussions of the actual chemiluminescent reactions assume that the emitting molecule is the initial hydrazide itself or some closely related compound.

It is a matter of practical and perhaps of theoretical importance to know if the emission curves, as observed experimentally, represent the emitting molecule. Distortion could occur by the presence of compounds which either absorb some of the emitted light or fluoresce under its influence. According to the published curves,[25, 27, 29, 30] absorption by the hydrazide itself should not greatly influence the emission in the visible region, although the same may not be true for other reactants and catalysts which have been used. That fluorescence of various dyes occurs when they are dissolved in a solution of the hydrazide emitting chemiluminescence, has been shown by a number of authors.[31-33] Also, mixtures of hydrazide with dimethylbiacridinium salt, in solutions where only the hydrazide would show chemiluminescence, give the green light of the acridinium chemiluminescence or fluorescence.[34] Tamamusi[35] has favored the idea that such light emissions may not be true fluorescences but actual transfers of energy by collision of hydrazide with dye followed by light emission from the dye. A transfer of energy during collision of sodium chloride and atomic sodium is included in the gas reaction described above. However, no similar process has been shown for chemiluminescence in solution and, according to Weber and Ochsenfeld,[36] the experimental results on hydrazide can be quantitatively explained without assuming such an energy transfer.

The relation of structure to the luminescence has been clarified by work from a number of laboratories.[3, 9, 10, 37-45] Several conclusions were summarized by Drew[41] in the 1939 Faraday Symposium on luminescence, although agreement is not complete.[42, 43] According to Drew, the 6-membered ring hydrazide is an essential requirement for chemiluminescence. Neither the open chain nor the 5-membered ring hydrazides show any luminescence. Apparent exceptions to this rule are explained as due either to a preliminary change into the 6-membered ring compound or to the presence of traces of it as an impurity. It has also been found[9, 39, 40] that methyl group substitutions such as



give non-luminescent compounds. Therefore, it is concluded that both hydrogens must be present on the nitrogen atoms or as the di-enol. The sensitivity of the eye to light requires exceptionally rigorous purification of the compounds studied to avoid misinterpreting the data. Drew and Garwood[40] used as many as eight recrystallizations to free some materials of traces of chemiluminescent impurities. More recently, Huntress and Gladding[45] have shown that changing the order of the carbon atoms and nitrogen atoms in the ring in the several possible ureas and quinoxalines, results in compounds which do not luminesce although they appear to oxidize.

Given the 6-membered ring hydrazide, Albrecht[8] showed in his thesis that substituents in the other ring greatly influenced the amount of light emitted. An amino group in the 3 or 4 position increased the amount of light as compared with the unsubstituted phthalhydrazide. Position 3 was more effective than 4. The nitro group decreased the amount of light. Drew[41] concluded that, in general, substitutions in the *ortho* positions with respect to the cyclohydrazide ring had greater quantitative effects than the corresponding substitutions in the *meta* positions. The nature of the substituent in the benzene ring has a great effect on the intensity of the light emitted and also influences its color.[8, 38] Some relative intensity measurements have been made[38, 40, 42] which permit a listing of substituent groups in the approximate order of light emitted. Groups such as NH_2, NHMe, OH, $NH \bullet NH_2$, in that order, and to a lesser extent chlorine, bromine, and iodine, increase the amount of light emitted as compared to the unsubstituted compound. On the other hand, groups such as NO_2 increase the light very little or decrease it. In general, Drew and Pearman[38] consider the first class to correspond to the *ortho* and *para* directing groups of organic chemistry, and the second to *meta* directing groups.

It may be that this apparent relation between the amount of chemiluminescence and the directive influence of substituent groups comes about through the usual effect of the groups on the chemical reaction. In this connection, however, some results obtained by West[46] on the fluorescence of naphthalene and its derivatives are of interest. The following relative efficiencies of fluorescence are reported:

NH_2 : 1
β-OH : 0.75
β-CN : 0.5
α-COOH : 0.3
Naphthalene : 0.15
α-chlor : 0.05
NO_2 : none.

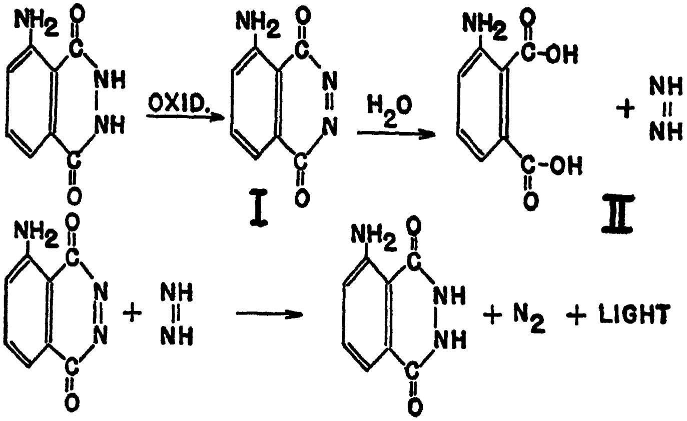
Although these results presumably do not involve a chemical reaction, the various groups again have a striking and specific effect on the quantitative result. Also, the order of the groups is approximately the same as for the chemiluminescence, with the exception of cyanide and carboxyl which are not given by Drew and Pearman.[38] These two are *meta* orienting groups.

Compounds present in solution also have a marked influence on the amount of chemiluminescence, just as in fluorescence. For instance, it has been found that traces of hydroquinone[24, 42] cause a great diminution in the light intensity. A series of papers by Weber and his collaborators have reported studies on the influence of many compounds and the halide ions on the fluorescence[47] and chemiluminescence[21–23] of the hydrazide. The influence on chemiluminescence may operate both through the chemical reaction and through quenching, which is analogous to fluorescence quenching. That the effect of chloride and bromide ions on a chemiluminescent reaction need not be to decrease but may increase the amount of light emission, has been shown with the *Cypridina* luciferin and luciferase system.[48]

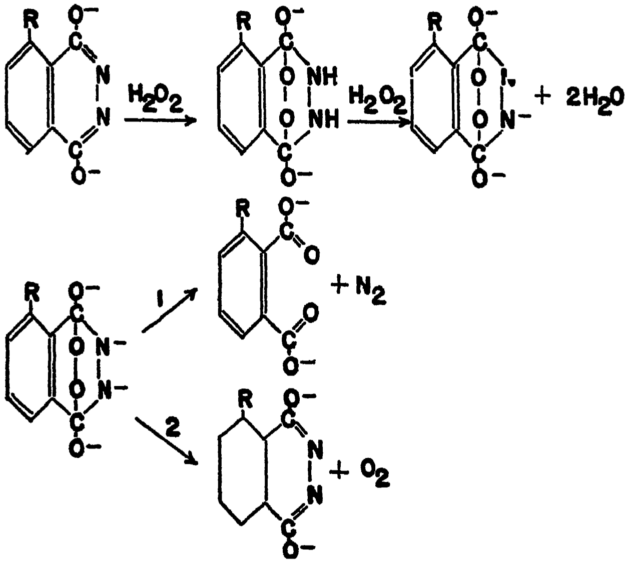
The actual reactions which produce the chemiluminescence are not agreed upon, although several authors[3, 21, 24, 41, 49] have suggested systems which account for a certain amount of the experimental data. Albrecht's[8] original formulation (see following page) has not been accepted.[41]
The azodiacyl compound (I) is unknown, and the diimine (II) is hypothetical. However, in 1942, the finding of a compound, colored and unstable, believed to be (I) was reported by Kautsky and Kaiser.[50] A solution of this material, free of oxidizing agent or oxygen, is said to emit light when made alkaline. No further work on its isolation and identification has appeared.

Meanwhile, Drew and Garwood[40] reported the isolation of a sodium salt of the peroxide of the 3-aminophthalcyclohydrazide. This compound, dissolved in water, gives off luminescence when hemoglobin is added to the solution. The intensity increases when alkali is added.

The same authors also report that the extent of the destruction of the hydrazide during chemiluminescence depends on the conditions, includ-



ing particularly the oxidizing agent used. Destruction, with evolution of all of the hydrazide and much of the amino nitrogen, occurs with hypochlorite and permanganate. However, when to a solution of hydrazide and hydrogen peroxide, sodium hypochlorite was added slowly, as little as one-seventh of the hydrazide nitrogen was liberated. The authors suggest, therefore, that most of the hydrazide can be recovered unchanged and that the destruction of this ring may be a side reaction and not an integral part of the chemiluminescent reaction. Their provisional formulation is as follows, on the assumption that the hydrazide in alkaline solution is in the enol form and ionized although no titration curve is presented:



Reaction 1 leads to the release of nitrogen and need not be luminescent. Reaction 2 may be the major source of luminescence. This, unlike most previous ideas, suggests that the hydrazide is not the substrate but a catalyst for the decomposition of hydrogen peroxide, and it has been so considered in some kinetic studies.[25] A study of the reactions of the isolated hydrazide peroxide, including quantitative determinations of nitrogen and oxygen, should be revealing.

Stross and Branch,[24] also in 1938, on the basis of quantitative studies on the reactants, ferricyanide and hydrogen peroxide, and accounting for the oxygen, concluded that the chemiluminescence was associated with a two-step oxidation, probably to the azo compound postulated by Albrecht.[3] In contradiction to Albrecht and in agreement with Drew,[41] they consider that the luminescence comes before the breaking of the hydrazide ring. They consider that in their reaction system the ferricyanide produces a one-unit oxidation to a free radical which then reacts with hydrogen peroxide.

Additional proposals and experiments have been contributed by others.[21, 25, 49] Weiss[49] combines a number of chemiluminescences into a single scheme. Lophine, after hydrolysis, luciferin, and the hydrazide have fairly labile hydrogen atoms which, he states, can be removed by oxidation. In general, Weiss says that the reaction can be written, $AH_2 \rightarrow AH \rightarrow A$, where A represents all of the compound except the two labile hydrogen atoms. This formulation also has the free radical, AH, as an intermediate step. The actual luminescence arises, according to Weiss's theory, by combination of radicals or ions:

$$2\,AH \rightarrow AH_2 + A^*$$

$$\text{or, } AH \rightleftarrows A^- + H^+ \text{ and } AH + A^- \rightleftarrows AH^- + A^*.$$

Weber and his collaborators[21] include both the peroxide of Drew[40] and Albrecht's[3] system in a combined series of reactions.

It is evident that the problem has been left in a confused state, perhaps because of the war. Most authors, except Drew, include the azo compound to which Kautsky has suggested that he has a direct approach, but for which he has published only fragmentary evidence. Drew's isolation of the peroxide, on the other hand, seems definitive, but limited qualitative and quantitative data have been published upon its relation to the luminescent reaction, action of catalysts, etc. All formulations seem highly provisional until these basic matters and the fundamental position of the hydrazide in the luminescent reaction are settled.

The Biacridinium Salts

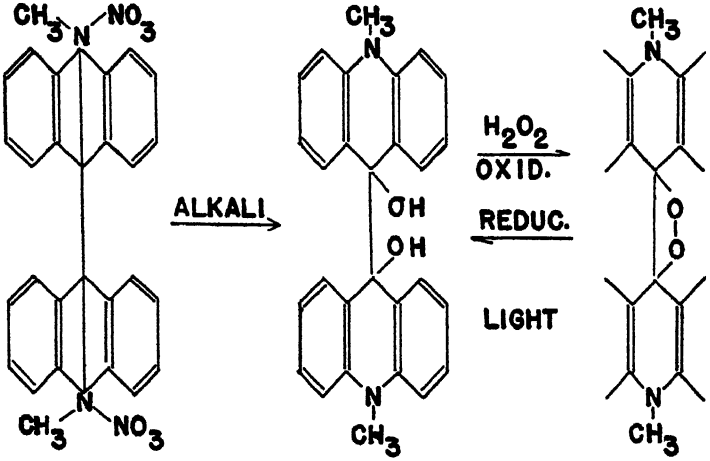
Gleu and Petsch,[4] in 1935, showed that N,N′-dimethylbiacridinium nitrate* gives a chemiluminescence when treated with hydrogen peroxide in alkaline solution. The chemiluminescence is green, as is the fluorescence. The compound is yellow. The intensity of the light emitted varies widely, depending upon the conditions. In solutions made alkaline with sodium hydroxide, the luminescence is relatively dim when hydrogen peroxide is added, but it lasts for a long time. The intensity can be increased enormously by the addition of osmium tetroxide as a catalyst and the time of luminescence decreased to one second. Under such conditions, light could be observed at a concentration of luminescent substance of 10^{-10} M. When the reaction is carried out in a medium of concentrated ammonium hydroxide, the light with hydrogen peroxide alone is much brighter than in a sodium hydroxide solution. It may then last for only a minute. Osmium tetroxide has less effect on the reaction in concentrated ammonia. The quantitative effects of temperature,[51] the kinetics of the reaction,[25, 26] and the quenching or enhancing effects of foreign substances[36, 52] have been studied.

---

* Also called lucigenin.

Unexpectedly, when Gleu and Petsch tried other oxidizing agents, they observed no luminescence. However, they found that a number of common reducing agents, such as hydrosulfite or stannite, produced a chemiluminescence if oxygen was also present. None was observed if the solution was first freed of oxygen by passing nitrogen through it. In view of the need for oxygen in order to obtain luminescence, the authors concluded that the reaction was both an oxidation and a reduction. Since hydrogen peroxide is both an oxidizing and a reducing agent, this was not inconsistent with the other findings.

A tentative formulation designed to explain these several facts was given as follows:



According to this idea, the biacridinium salt would function essentially as a catalyst. With hydrogen peroxide, the net reaction would be:

$$2\,H_2O_2 \rightarrow 2H_2O + O_2 + 46\text{ Cals.},$$

and with the other reductants, the oxidation of the reducing agent by atmospheric oxygen. One difficulty immediately arises, since the 46 Cals. are about 10 Cals. too small for some of the wavelengths represented in the light emission. However, Gleu considers that this does not definitely prove the reaction to be inadequate.

Comparisons of the spectral distribution curves of fluorescence and chemiluminescence have also been made for the biacridinium salts. The compound fluoresces in water and in acid solution. According to Weber, it will also fluoresce in mildly alkaline sodium carbonate solutions but, like the hydrazide, the chemiluminescence and fluorescence appear to have widely different pH optima. The original dimethylbiacridinium compound was studied by Eymers and Van Schouwenberg,[28] who concluded that the fluorescent and chemiluminescent spectra were identical, even though the media in which they were studied were not identical. These same authors have extended their work by comparing fluorescent and chemiluminescent spectral distribution curves, including bioluminescent sources, after analysis. The analysis was based on the assumption that the observed graph, when intensity was plotted against frequency, was made up of two or more essentially symmetrical curves. The frequencies of these postulated fundamental curves tended to be the same for different reactions. How much the results depend on the above described secondary fluorescence is unknown.

More recently, however, Gleu and collaborators[53-55] have prepared a number of biacridinium derivatives and compared the color of the fluorescence and chemiluminescence.[55] The diethyl derivative shows identical green fluorescence and chemiluminescence. The same is true for the diphenyl at some concentrations. However, below 10^{-6} M, the chemiluminescence becomes blue. Fluorescence and chemiluminescence also differ in other compounds. The authors conclude, therefore, that the supposed agreement, which they and others had previously reported and used as a partial basis for the above formulation, is merely a coincidence which occurs with some compounds.

Kautsky and Kaiser,[50] in 1943, reported that all previous interpretations of the spectral distribution curve of the chemiluminescence are faulty. As ordinarily obtained, they consider the light emission to be a mixture of a true chemiluminescence and the green fluorescence of the biacridinium nitrate. They state that when the reaction is carried out in a dilute solution at 40—50° C., the fluorescence is largely absent and the true chemiluminescence, blue in color, is obtained. This seems to fit the observations of Gleu and Schaarschmidt[55] mentioned above. Further, Kautsky and Kaiser find that the spectral distribution curve of this "true chemiluminescence" is closely similar to the fluorescence of N-methylacridone under exactly the same conditions. It is the acridone, therefore, which these authors consider as being the primary emitter of the luminescence, and they regard its formation from the carbinol with liberation of 65 Cals. as the chemiluminescent reaction. If this is true, there should be a striking effect of pH on the color of the emission curve as usually obtained, since the fluorescence of the biacridinium salt is reported as decreasing in alkaline solution. Also, the question arises as to how the reducing agents produce their effect, since the chemiluminescent reaction, as formulated with hydrogen peroxide by these authors, is an oxidation.

Unfortunately, no additional work has been found, so that here, also, the fundamental role in the luminescent reaction, substrate or catalyst, of the major compound is in doubt. The formation of acridone as a possible side reaction was mentioned by Gleu and Petsch.[4]

BIBLIOGRAPHY

1. **Evans, M. G., H. Eyring, & J. F. Kincaid**
   1938. J. Chem. Phys. **6**: 349.
2. **Audubert, R.**
   1938. Angew. Chem. **51**: 153.
   1939. Trans. Faraday Soc. **35**: 197.
3. **Albrecht, H. O.**
   1928. Dissertation, Kaiser Wilhelm-Institut.
   1928. Z. Phys. Chem. **135**: 321.
4. **Gleu, K., & W. Petsch**
   1935. Angew. Chem. **48**: 57.
5. **Polanyi, M., & G. Shay**
   1928. Z. Phys. Chem. **B1**: 30.
6. **Harvey, E. N.**
   1929. J. Phys. Chem. **33**: 1456.
   1940. Living Light: 113–121. Princeton University Press.
7. **Huntress, E. H., L. N. Stanley, & A. S. Parker**
   1934. J. Chem. Ed. **11**: 142.
   1934. J. Am. Chem. Soc. **56**: 241.
8. **Witte, A. A. M.**
   1935. Rec. Trav. Chim. **54**: 471.
9. **Gleu, K., & K. Pfannstiel**
   1936. J. Prakt. Chem. **146**: 137.
10. **Wegler, R.**
    1937. J. Prakt. Chem. **148**: 135.
11. **Tamamusi, B.**
    1937. Naturwiss. **25**: 318.
12. **Thielert, H., & P. Pfeiffer**
    1938. Berichte **71B**: 1399.
13. **Cook, A. H.**
    1938. J. Chem. Soc.: 1845.
14. **Schales, O.**
    1938. Berichte **71B**: 447.
    1939. Berichte **72B**: 167.
15. **Vasserman, E. S.**
    1939. C. R. Acad. Sci. U.R.S.S. **24**: 704.
16. **Briner, E.**
    1940. Helv. Chim. Acta **23**: 320.

17. **Schneider, E.**
1941. J. Am. Chem. Soc. **63**: 1477.
18. **Steigmann, A.**
1941. Chem. & Ind. **889.**
19. **Geyer, B. P., & H. M. Haendler (with G. McP. Smith)**
1941. J. Am. Chem. Soc. **63**: 3071.
20. **Geyer, B. P., & G. McP. Smith**
1942. J. Am. Chem. Soc. **64**: 1649.
21. **Weber, K., A. Rezek, & V. Vouk**
1942. Berichte **75B**: 1141.
22. **Weber, K., & M. Krajcinovic**
1942. Berichte **75B**: 2051.
23. **Weber, K., W. Lahm, & E. Hieber**
1943. Berichte **76B**: 366.
24. **Stross, F. H., & G. E. K. Branch**
1938. J. Org. Chem. **3**: 385.
25. **Sveshnikov, B. Ya.**
1938. Acta Physicochim. U.R.S.S. **8**: 441.
1942. C. R. Acad. Sci. U.R.S.S. **35**: 278.
26. **Sveshnikov, B. Ya., & P. P. Dickun**
1942. Acta Physicochim. U.R.S.S. **17**: 173.
27. **Harris, L., & A. S. Parker**
1935. J. Am. Chem. Soc. **57**: 1939.
28. **Eymers, J. G., & K. L. van Schouwenberg**
1936. Enzymologia **1**: 107.
1937. Enzymologia **3**: 235.
29. **Briner, E., & E. Perrottet**
1940. Helv. Chim. Acta **23**: 1253.
30. **Zelinskii, V. V., & B. Ya. Sveshnikov**
1942. C. R. Acad. Sci. U.R.S.S. **34**: 252.
31. **Plotnikov, I., & J. Kubal**
1938. Phot. Korr. **74**: 97.
32. **Kubal, J.**
1938. Phot. Korr. **74**: 132.
33. **Plotnikov, I., M. Doljak, & T. Kopsić**
1940. Phot. Korr. **76**: 43.
34. **Schales, O.**
1939. Berichte **72B**: 1155.
35. **Tamamusi, B.**
1940. Naturwiss. **28**: 722.
36. **Weber, K., & W. Ochsenfeld**
1942. Z. Phys. Chem. **B51**: 63.
37. **Drew, H. D. K., & H. H. Hatt**
1937. J. Chem. Soc.: 16.
38. **Drew, H. D. K., & F. H. Pearman**
1937. J. Chem. Soc.: 26.
1937. J. Chem. Soc.: 586.
39. **Drew, H. D. K., H. H. Hatt, & F. A. Hobart**
1937. J. Chem. Soc.: 33.
40. **Drew, H. D. K., & R. F. Garwood**
1937. J. Chem. Soc.: 1841.
1938. J. Chem. Soc.: 791.
1939. J. Chem. Soc.: 836.
41. **Drew, H. D. K.**
1939. Trans. Faraday Soc. **35**: 207.

42. **Zellner, C. N., & G. Dougherty**
1937. J. Am. Chem. Soc. **59**: 2580.

43. **Vasserman, E. S., & G. P. Miklukhin**
1939. J. Gen. Chem. U.S.S.R. **9**: 606.
1940. J. Gen. Chem. U.S.S.R. **10**: 202.

44. **Huntress, E. H., & W. M. Hearon**
1942. J. Am. Chem. Soc. **64**: 86.

45. **Huntress, E. H., & J. V. K. Gladding**
1942. J. Am. Chem. Soc. **64**: 2644.

46. **West, W.**
1941. Ann. N. Y. Acad. Sci. **41**: 203.

47. **Weber, K.**
1942. Berichte **75B**: 565.

48. **Anderson, R. S.**
1937. J. Am. Chem. Soc. **59**: 2115.

49. **Weiss, J.**
1939. Trans. Faraday Soc. **35**: 219.

50. **Kautsky, H., & K. H. Kaiser**
1942. Naturwiss. **30**: 148.
1943. Naturwiss. **31**: 505.

51. **Tamamusi, B., & H. Akiyama**
1939. Trans. Faraday Soc. **35**: 491.

52. **Weber, K.**
1941. Z. Phys. Chem. **B50**: 100.

53. **Gleu, K., & S. Nitzsche**
1939. J. Prakt. Chem. **153**: 200.
1939. J. Prakt. Chem. **153**: 233.

54. **Gleu, K., & A. Schubert**
1940. Berichte **73B**: 805.

55. **Gleu, K., & R. Schaarschmidt**
1940. Berichte **73B**: 909.

THE CHEMISTRY OF *CYPRIDINA* LUCIFERIN

By AURIN M. CHASE

Physiological Laboratory, Princeton University, Princeton, N. J.

Luminescence is encountered throughout the animal kingdom and in certain lower forms of plant life. However, the luminescent reaction has been demonstrated *in vitro* in only five orders of animals. These include certain beetles (fireflies), one mollusc (*Pholas*), certain ostracods (*e.g., Cypridina*), a few worms (*e.g., Odontosyllis*), and at least one decapod, a deep-sea form. The most favorable organism as a source of material for the study of the luminescent reaction *in vitro* is the ostracod crustacean, *Cypridina hilgendorfii,* whose eminent suitability for the extraction of the compounds concerned was pointed out by Harvey (1917). This animal possesses a gland in the head region, where luciferase, the enzyme which catalyzes the reaction, and luciferin, the substrate, are produced. When alive, the animal can eject these two compounds, in the form of two kinds of granules, into the sea water through separate pores, and a blue luminescence then occurs in the water. If the organisms are dried immediately after being caught, and are kept dry, the luciferin and luciferase in the gland remain stable for years and can be extracted when desired. The paper by Harvey (1948) contains photographs of *Cypridina* and drawings of the gland, on page 335.

The simplest method of extracting luciferin from *Cypridina* is by grinding up the dry organisms and adding hot water. This inactivates the luciferase, so that no luminescence occurs in the extract but the luciferin is obtained unaltered. The suspension is immediately cooled to retard destruction of the luciferin by oxidation. After filtering, the clear solution contains luciferin and all the other water-soluble components of the organism. Such crude luciferin extracts, though easy to prepare, are not very suitable for chemical work, because of the great instability of the luciferin in the presence of dissolved oxygen and of oxidizing systems that are extracted along with the luciferin. Also, the impurity of such preparations is undesirable from a chemical point of view and, indeed, luciferin in crude extracts often behaves quite differently than does more highly purified luciferin.

Several methods of extracting and purifying *Cypridina* luciferin have been developed. Kanda (1924, 1929) extracted the dry, powdered organisms with methyl alcohol in absence of oxygen and, after various precipitations and re-solutions, obtained luciferin which was more stable against oxidation and certainly purer than that in a crude aqueous extract. The chemical work on luciferin prior to about 1940 is discussed in publications of Harvey (1940, 1941).

The greatest advance toward the purification of luciferin is due to the work of Anderson. To determine quantitatively the concentration of luciferin present in his various purification steps, he developed a photoelectric method for measuring the total light emitted by a luciferin solution. This apparatus (Anderson, 1933) does not measure light intensity directly, as do most photoelectric and all visual methods, but instead it yields the integral curve for the luminescent reaction. The output of the photoelectric cell is stored in a condenser and the charge accumulated on the condenser, after any time from the start of the reaction, can be balanced with a potentiometer, using a Lindemann electrometer. The method is therefore a null-point one, and light-emission can be measured in terms of millivolts, to about one millivolt, the limit set by the Lindemann electrometer. The capacity of the condenser determines the sensitivity of the method. With this apparatus, very weak luminescences can be measured quantitatively. A good example of the sort of data obtainable is given in FIGURE 2. The value of the ordinate at any time is a relative measure of the amount of luciferin that has reacted with luciferase to give luminescence, from the start of the reaction to that time. The slope of the curve at any time is proportional to the intensity of the luminescence at that particular time. The data secured by the method are, therefore, analogous to those obtained in any reaction where an end product is determined quantitatively at various times during the course of the reaction.

Anderson (1935) worked out a purification procedure which yields from *Cypridina* a luciferin of much greater purity than had been obtained previously. Briefly, the method is as follows. Dry, powdered organisms are extracted for twenty-four hours with absolute methyl alcohol that is kept free of dissolved oxygen by saturation with purified hydrogen in a special extraction vessel. The vessel is then opened and a small amount of *n*-butyl alcohol is added and the methyl alcohol is removed by evaporation at low pressure. This solution, after having been chilled, is treated with benzoyl chloride. The resulting derivative of luciferin not only does not give light on addition of luciferase, but is much more stable in the presence of air than is the luciferin in its original state. Water is now added to hydrolyze the excess benzoyl chloride, and the butyl alcohol fraction is dissolved in 10 volumes of water. The resulting solution is extracted with ether, into which most of the butyl alcohol and luciferin derivative pass, leaving highly colored impurities in the aqueous phase. The ether is next removed *in vacuo*, leaving the inactive luciferin, already considerably free of colored impurities, in the residual butyl alcohol. This solution is mixed with a large volume of 0.5 *N* HCl, saturated with hydrogen, and heated in a hydrogen atmosphere for an hour at 95°-100° C and then cooled in an ice water bath. The mixture, now containing active luciferin, is then washed again with ether. At this stage, the "reactivated" luciferin remains almost entirely in the aqueous phase, and considerable colored material passes into the ether phase. The luciferin is finally extracted from the aqueous solution with *n*-butyl alcohol, deaerated with hydrogen and, if desired, is put through the same cycle of purification a second time. Some yellow color remains even after three cycles of purification and, as will be apparent presently, this is undoubtedly a property of luciferin.

This method results in a purification of the luciferin of about 2,000 times, in terms of amount of light per unit of dry weight of solid material in the final solution as compared with the starting material. The stability of the luciferin against oxidation is also greatly increased, perhaps because oxidizing systems that were present in the original material have been removed during the purification procedure.

In most of the recent work which has been done on the chemistry of luciferin, the luciferin has first been subjected to Anderson's (1935) purification procedure. It has frequently been found that quite different experimental results are obtained when using this purified material than with crude aqueous extracts of *Cypridina.* For example, luciferin in crude extracts is not affected by even high concentrations of cyanide, whereas the more purified compound loses its light-producing characteristics in the presence of extremely low cyanide concentrations. (Giese and Chase, 1940.)

The luciferase used with this purified luciferin, when studying the luminescent reaction *in vitro,* is ordinarily prepared by dialyzing a water extract of powdered dry *Cypridina* against distilled water for some days at a low temperature until the solution has become practically colorless. Considerable inactive protein is precipitated by this treatment and most of the luciferin and other dialyzable compounds are certainly eliminated. That luciferin can pass through a dialysis membrane was demonstrated by Harvey (1917), and it seems reasonable to assume that oxidized luciferin is also dialyzable, although this has not actually been shown.

Anderson found (1936), with luciferin purified by his method (1935), that the luminescent reaction of luciferin and luciferase, which requires the presence of oxygen, was irreversible. On the other hand, he found that the non-luminescent oxidation of luciferin which occurs in the presence of dissolved oxygen or of certain oxidants (*e.g.,* ferricyanide) can be almost completely reversed by $Na_2S_2O_4$ or suitable reductants, if the latter are added soon enough. Anderson concluded that the luminescent reaction of luciferin and luciferase and the non-luminescent oxidation of luciferin are quite different from each other. He found a redox potential for the reversible oxidation of luciferin, assuming that the oxidation removed two electrons per molecule, about 0.01 volt negative to quinhydrone at 23° and pH 6.8 (Anderson, 1936). Korr (1936) found a similar value. E_0' is, then, about +0.260 volt at pH 7. Anderson pointed out the similarity of this value to the redox potentials encountered in the case of certain naturally occurring polyhydroxybenzene derivatives studied by Ball and Chen (1933). He has emphasized the possibility that such a grouping may be present as part of the luciferin molecule. It is interesting to note that, in the work of Ball and Chen, the oxidized forms of the natural compounds whose potentials they measured were very unstable. In fact, it was necessary to use a flow technique in order to make the measurements. This feature of these compounds again shows a similarity to the luciferin-oxidized luciferin system, where the oxidized luciferin also seems to be unstable, since it cannot be reduced if it has stood too long in the presence of an oxidizing agent or of dissolved oxygen.

Anderson (1936) studied the reversible non-luminescent oxidation of purified luciferin in some detail and found that the luminescent reaction which occurred upon adding luciferase to a solution of luciferin that had been exposed to air for a short time showed evidence of two distinct kinds of luminescence, a rapid and a slow luminescence, occurring simultaneously. He attributed the intense, rapid luminescence to reduced luciferin in the solution and the dim, slow luminescence to reversibly oxidized luciferin which was being reduced in the presence of luciferase or of compounds extracted with it.

The oxidation of luciferin can be accelerated by irradiation with certain parts of the spectrum, as was shown by Harvey (1925, 1926). He used crude luciferin extracts from *Cypridina* and found that blue, violet, and near ultraviolet light were all effective in accelerating the oxidation. Chase and Giese (1940) later found that, when purified luciferin was used, the ultraviolet below 3000 Å was effective, but ultraviolet of longer wavelength and the visible part of the spectrum were not, unless a photo-sensitizing compound was added. A number of compounds were found to act as sensitizers, including eosin and fluorescein. Boiled, crude aqueous extracts of *Cypridina,* containing oxidized luciferin and all the other water-soluble components of the organism, were very effective, indicating that impurities in the crude extract were sensitizing the oxidation of luciferin irradiated by visible light in the case of Harvey's earlier experiments.

Chase and Giese found that, when the luminescent reaction was measured upon adding luciferase to luciferin that had been irradiated with ultraviolet light or with visible light in the presence of a sensitizer, the reaction could be differentiated into two parts, one a bright luminescence and the other a dim luminescence, as Anderson had found in the case of luciferin that had stood in the presence of dissolved oxygen. An interesting exception occurred when the purified luciferin was irradiated with visible light in the presence of riboflavin as a sensitizer. In this case, the luminescent reaction which ensued on addition of luciferase lacked the dim component, as shown in FIGURE 1. Riboflavin evidently prevents the reduction by luciferase (or whatever is responsible) of reversibly-oxidized luciferin. However, riboflavin is effective in this way only if the oxidation of luciferin is caused by the action of light with the flavin as sensitizer. If the luciferin is oxidized by dissolved oxygen, in the dark, with riboflavin present in the solution, and luciferase is then added, the resulting luminescence shows the bright and dim component ordinarily observed.

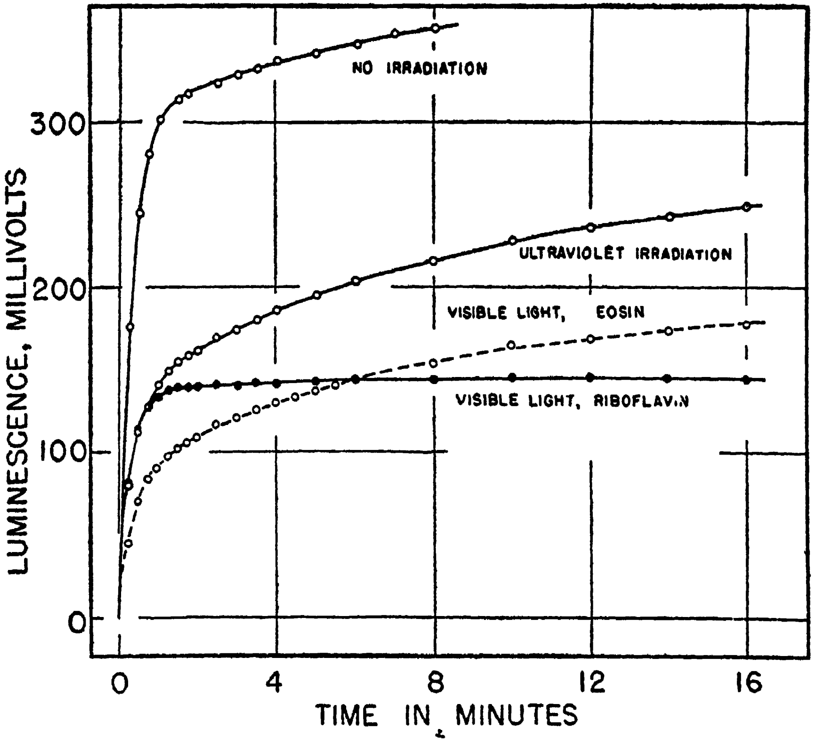


FIGURE 1. Typical luminescence curve obtained on adding luciferase to luciferin solutions which have been irradiated for three minutes with visible light in the presence of about 0.001 per cent concentration of riboflavin. Although an initial bright luminescence occurs, there is no light-emission after two minutes. The luminescence curves obtained after irradiation of luciferin with visible light with eosin present, and after ultraviolet irradiation without a sensitizer, both show not only a bright luminescence during the first two minutes, but also a dim luminescence which persists for at least sixteen minutes.

Recently, Chase and Lorenz (1945) separated the velocity constants of the enzyme-catalyzed luminescent reaction of luciferin and of its non-luminescent oxidation. They measured the luminescent reaction at five temperatures and fitted the data with an equation representing two first-order reactions occurring simultaneously. FIGURE 2 shows their experi-

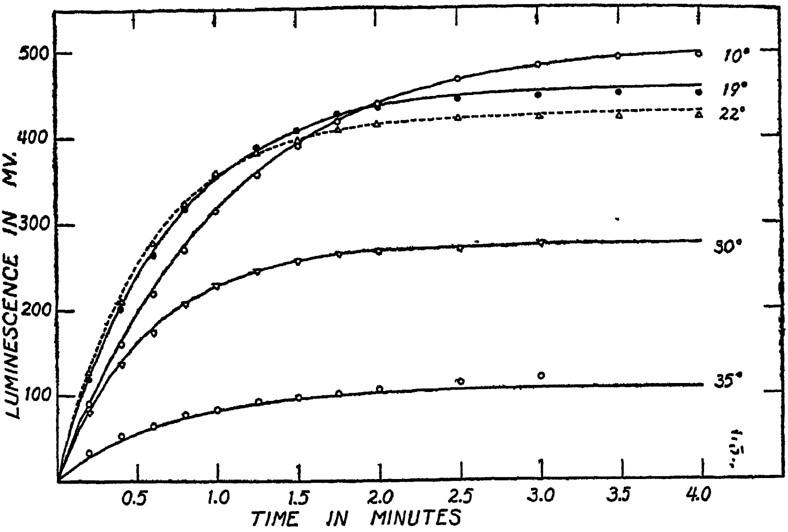


FIGURE 2. The luminescent reaction at five temperatures. The symbols represent experimentally measured total light emitted from the moment of mixing luciferin and luciferase to various times. The curves are theoretical, calculated from an equation derived on the assumption that two first order reactions are occurring simultaneously: one, the luminescent reaction of luciferin, and the other, its non-luminescent oxidation.

mental points and the curves which describe them are theoretical. Since Anderson (1936) has demonstrated that the luminescent reaction of luciferin and luciferase is a different process from the non-luminescent oxidation of luciferin, it is clear that both these processes may occur simultaneously. Chase and Lorenz assumed that, since dissolved oxygen is probably present in excess, the non-luminescent oxidation of luciferin might obey a first-order equation even though it might actually be a bimolecular reaction. Amberson had shown (1922) that the luminescent reaction obtained on mixing crude *Cypridina* luciferin and luciferase extracts obeyed the equation for a first order reaction. A record of the logarithmic decay of *Cypridina* luminescence is shown in the paper by Harvey (1948). The equation which describes the curves of FIGURE 1 was derived by Chase and Lorenz on the assumption, then, that the luciferin is the principal substrate in two first-order reactions occurring simultaneously, only one of which results in light emission. If x is the luciferin consumed by the luminescent reaction after time, t, and y is the luciferin consumed by the non-luminescent oxidation reaction after time, t, and if k_1 and k_2 are the respective velocity constants and a is the concentration of luciferin initially present, then:

$$\left\{\begin{aligned} \frac{dx}{dt} &= k_1(a-x-y) \\ \frac{dy}{dt} &= k_2(a-x-y) \end{aligned}\right\}\text{are to hold simultaneously.}$$

This gives $\frac{dy}{k_2} = \frac{dx}{k_1}$, whence, since $y = 0$ when $x = 0$, $y = \frac{k_2}{k_1}x$.

Using this value of y, $\frac{dx}{dt} = k_1\left(a - x - \frac{k_2}{k_1}x\right)$.

On integration and simplification, one obtains

$$x = \frac{k_1 a}{k_1 + k_2}\left(1 - e^{-(k_1 + k_2)t}\right)$$

This equation describes the data of FIGURE 2 rather exactly, the values of k_1 and k_2 being obtained directly from the experimental measurements.

If the logarithm of k_2 (the velocity constant of the non-luminescent oxidation reaction) be plotted against the reciprocal of the absolute temperature, a fairly straight line is obtained, as shown in FIGURE 3, and the slope of this line corresponds to an energy of activation of about 25,000 calories per mole. The velocity constant of the luminescent reaction, on the other hand, has a maximum value at about 23° C, the principal effect of temperature in this case being upon the enzyme, luciferase.

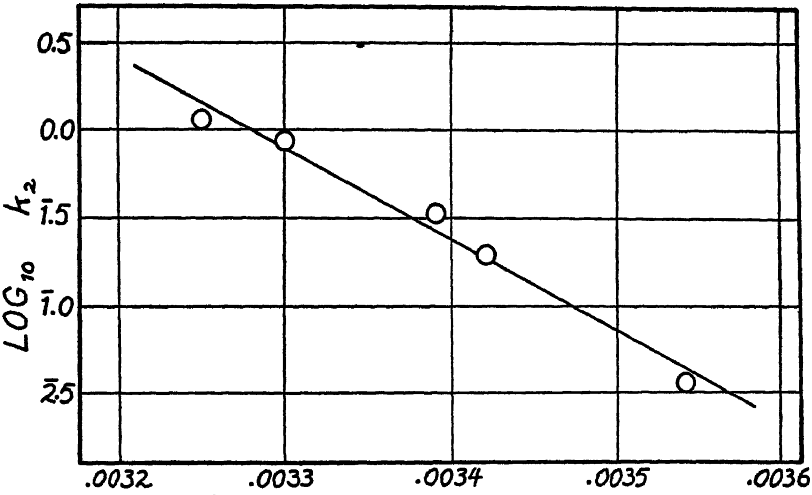


FIGURE 3. Logarithm₁₀ of the calculated velocity constant of the non-luminescent oxidation of luciferin plotted against the reciprocal of the absolute temperature. The slope represents a value of about 25,000 calories for the energy of activation.

Although the subject of this paper is the chemistry of luciferin, rather than the kinetics of the luminescent reaction, it seems proper to include one more datum from the kinetic point of view, in addition to those already mentioned.

Chance, Harvey, Johnson, and Millikan (1940) measured, by means of a special technique, the luminescent reaction of unpurified *Cypridina* luciferin and luciferase mixed in two different ways. In the first method, luciferin, in solution containing dissolved oxygen, was mixed with luciferase, also in solution containing oxygen. In the second method, a de-aerated solution containing both luciferin and luciferase was mixed with a solution containing dissolved oxygen. The luminescent reaction was much faster when the luciferase and luciferin were previously mixed than it was in the other case. The data were interpreted as showing that a combination between luciferin and luciferase is an essential requirement for luminescence, and that this combination represents a relatively slow reaction. The analysis also showed that the enzyme, luciferase, is the light-emitting molecule. This latter interpretation has generally been made, although it has not been proved that the light-emitting molecule may not be luciferin.

So far as the chemical structure of *Cypridina* luciferin is concerned, very little is known with certainty. However, several hypotheses have been advanced, most of them backed by some experimental evidence. The method of purification rules out any possibility that luciferin is a protein, while the redox potential measurements mentioned above have shown a possible relationship to hydroxybenzene derivatives. Anderson (1936) also interpreted the formation of benzoyl derivatives of luciferin as indicating the presence of reactive hydrogen which might be attached to nitrogen, sulfur or oxygen. Chakravorty and Ballentine (1941), however, found no nitrogen or sulfur in luciferin.

Giese and Chase (1940) found that luciferin purified by Anderson's method, irreversibly lost its property of giving luminescence with luciferase when very small concentrations of cyanide were present, as illustrated in FIGURE 4. On the other hand, luciferin in crude extracts of *Cypridina* is not affected by high cyanide concentrations. This indicated a chemical reaction between luciferin and cyanide, and Giese and Chase calculated a molecular weight for luciferin of between 800 and 2400, assuming the luciferin to be 100 per cent pure and assuming a 1:1 combination between the luciferin and the cyanide. Since the luciferin is certainly not pure, the true value for the combining weight is probably at least as low as 800 or even less. Giese and Chase interpreted the reaction with cyanide as due to cyanhydrin formation, and assumed an aldehyde or keto group on the molecule, which group might also be the site of combination with luciferase in the irreversible luminescent reaction.

With a view toward obtaining corroborative data on the combining weight of luciferin, an attempt was made recently, in collaboration with Dr. C. B. Anfinsen (unpublished), to determine an oxygen uptake by purified luciferin during its non-luminescent and luminescent oxidations, using the Cartesian diver method, and assuming two atoms of oxygen

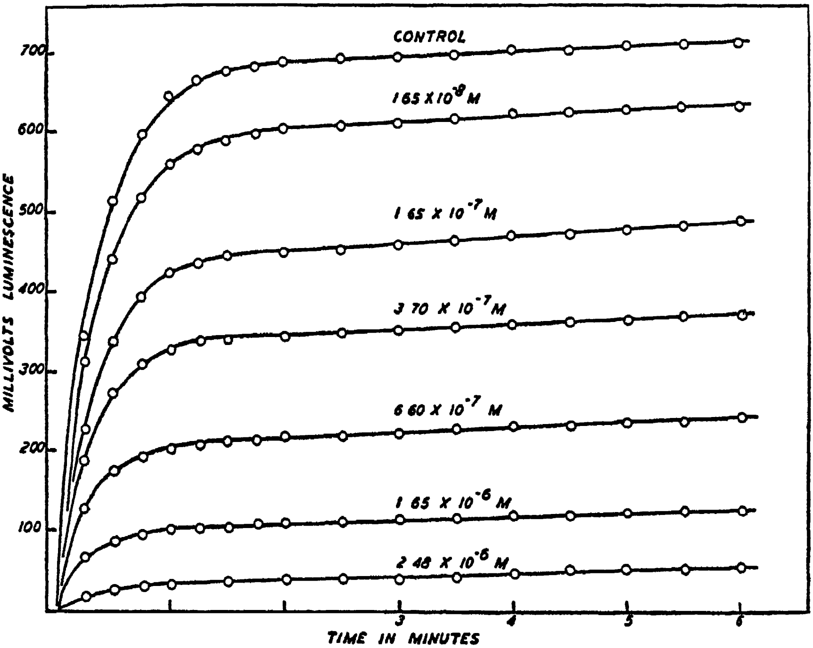


FIGURE 4. Quantity and rate of light-emission from identical samples of luciferin which were treated with various concentrations of cyanide, indicated on the curves, before addition of luciferase. Since the velocity constants of the curves are essentially the same, the luciferin rather than the luciferase is affected by the cyanide.

combined per molecule of luciferin in the non-luminescent oxidation. Although there were indications of oxygen uptake in some cases, the results were inconclusive. Parallel tests on the stability against oxidation of comparably small samples of luciferin solution, made it appear likely that the combination with oxygen, assuming that it does occur, took place during the time the divers were being filled and before the actual measurements of gas uptake could be started. R. S. Anderson has unpublished data on the concentration of ferricyanide necessary to oxidize known amounts of luciferin, which indicate a combining weight of about the same order of magnitude found in the case of the cyanide experiments (personal communication). The uncertainty in all measurements of the combining weight of luciferin by such methods lies chiefly in the assumption as to the purity of the luciferin. The cyanide data of Giese and Chase certainly point to a combining weight (and probably molecular weight) of less than 1000, perhaps about half this value. There seems little doubt that luciferin is a relatively small molecule.

Chase found (1942) that the luminescent reaction of purified luciferin and luciferase was reversibly inhibited by sodium azide. Since the total light was affected but not the velocity constant, the action of the azide was attributed to a reversible combination with the luciferin rather than with the luciferase. As was pointed out in another connection by Johnson, Eyring, and Williams (1942; see page 259), this interpretation may not be justified in the case of a reversible combination. The reaction might equally well take place between the azide and luciferase, although this does not seem likely. FIGURE 5A shows the effect of azide concentration upon the total light emitted. Experiments were run at pH 5.4 and at pH 6.6 and, as the figure shows, the azide was more effective at the lower pH. Since the pK of hydrazoic acid is about 4.7, a higher concentration of the undissociated acid would be present at pH 5.4 than at pH 6.6. For this reason, the effect was attributed to the undissociated acid rather than the azide anion.

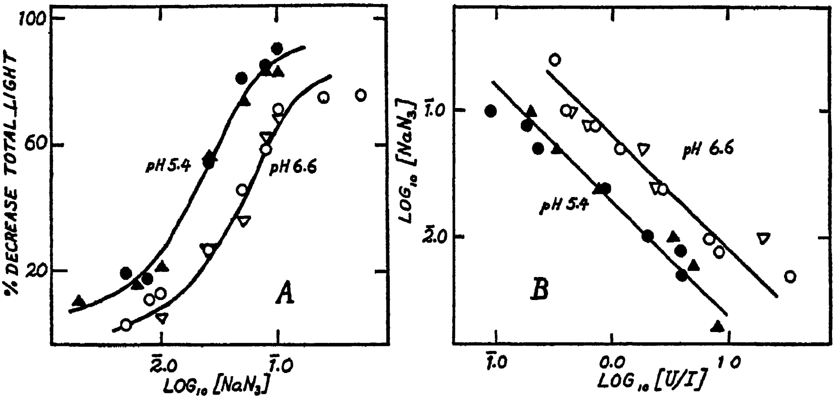


FIGURE 5. A. Per cent decrease of total light plotted against the logarithm of the sodium azide concentration. Four series of experiments are shown, two at pH 6.6 and two at pH 5.4. A given azide concentration is more effective at the lower pH, indicating that HN_3 may be involved.

B. The same data plotted in terms of a mass law equation. The slopes of the lines are approximately unity, which may indicate that one azide molecule or ion combines with a single luciferin molecule.

Since this reaction with azide is reversible, the data can be put into a form that makes them analyzable in terms of a mass law equation, as has been done in cases of inhibitor action on enzyme systems (see, for example, Fisher and Öhnell, 1940). In the present case, the analysis is as follows:

$$\text{Luciferin} + a \text{ azide} \rightleftarrows \text{luciferin (azide)}_a$$

$$\frac{[\text{Luciferin}]\,[\text{Azide}]^a}{[\text{Luciferin (Azide)}_a]} = K$$

$$\frac{[\text{Luciferin}]}{[\text{Luciferin (Azide)}_a]} = K \cdot [\text{Azide}]^{-a}$$

If [*Luciferin*] is called U and is put equal to the percentage of luminescence that occurs at any given azide concentration, taking as 100 per cent the luminescence obtained when no azide is present, and if [*Luciferin (Azide)$_a$*] is called I and is put equal to the percentage of potential luminescence that does not occur, then,

$$\frac{U}{I} = K \bullet [\text{Azide}]^{-a}, \text{ or } \log \frac{U}{I} = -a \log [\text{Azide}] + \log K.$$

Therefore, by plotting $\log_{10} \frac{U}{I}$ against the $\log_{10}$ of the azide concentration, a straight line should be obtained and the slope of the line should give the value of a, the number of azide molecules which combine with a single luciferin molecule. FIGURE 5B shows the data plotted this way. It is apparent that the slope is about –1, indicating one molecule of hydrazoic acid combining reversibly with each molecule of luciferin.

Fieser and Hartwell (1935) studied the reaction of hydrazoic acid with benzo- and naphthoquinones. According to them, an azido-hydroquinone is first formed, changing to an amino-quinone with liberation of nitrogen. Although a reaction of this sort would hardly be considered as readily reversible, it seems possible that a related type of reaction might occur in the case of luciferin and, if so, would be further evidence for a quinonoid structure.

Chase has shown (1945) that luciferin, freshly dissolved in pH 6.8 phosphate buffer in absence of air, has an absorption band at about 435 mμ in the visible spectrum. Measurements of the absorption spectrum were made with a recording spectrophotometer (Hardy, 1935). If dissolved oxygen is present, this band is rapidly replaced by another at about 465 mμ and this latter band then disappears slowly, leaving a practically colorless solution (Chase, 1943). These changes in the visible absorption spectrum of luciferin solutions take place much more slowly if the pH is more acid than 6.8. For example, as FIGURE 6 illustrates, at pH 5.1 the change is only about one-fifth as fast. Furthermore, at pH 5.1 there is definite evidence, from the absorption spectrum, for the production of a compound having acid-base indicator properties during the exposure of the luciferin to dissolved oxygen. This compound absorbs more strongly in the violet and near ultraviolet at a pH of 5.1 than it does at pH 6.8, as is also shown in the curves of FIGURE 6. The rate of change of the absorption spectrum increases with increase of pH in about the same way that the rate of non-luminescent oxidation of luciferin varies with pH in solutions containing dissolved oxygen (Chase, 1940). In fact, there is little doubt (Chase, 1943) that the amount of labile color in a luciferin solution is directly proportional to the concentration of luciferin, as measured by the total light obtainable from the solution.

If the visible absorption spectrum of a luciferin solution is measured during the luminescent reaction of luciferin and luciferase (as can be done with the Hardy recording spectrophotometer, which is not affected

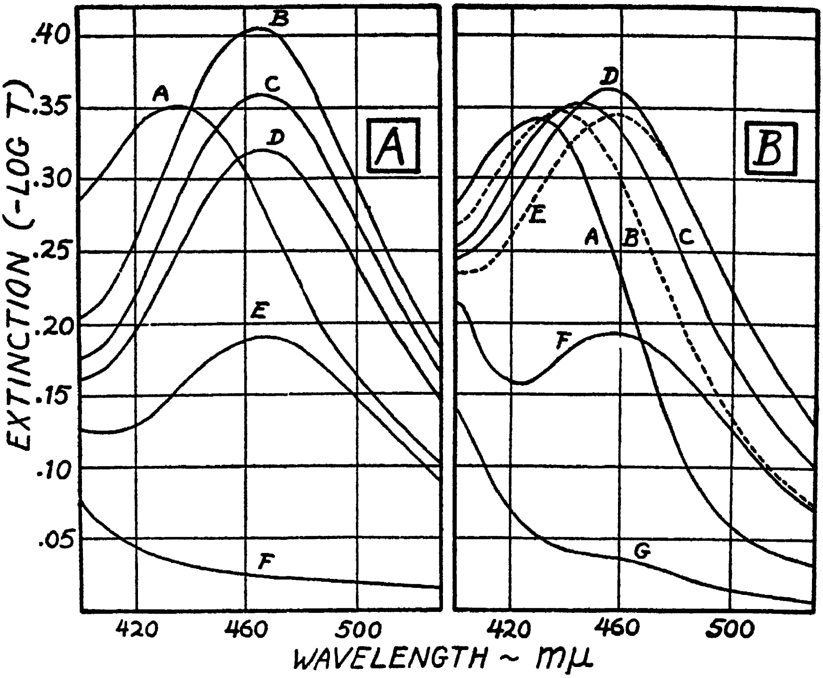


FIGURE 6. The left half shows the changes which occur in the visible absorption spectrum of a luciferin solution at pH 6.8 during exposure to air. Curve A is the spectrum 5 minutes after the luciferin was dissolved. Curve B was measured after a total elapsed time of 19 minutes, curve C after 31 minutes, curve D after 44 minutes, curve E after 81 minutes, and curve F after 2 days.

The right part of the figure shows similar measurements on a luciferin solution at pH 5.1. Curve A was measured 4 minutes after the luciferin was dissolved, curve B after 16 minutes, C after 30 minutes, D after 59 minutes, E after 81 minutes, F after 300 minutes, and G after 26 hours.

by continuous light-emission in the sample), it is found that the same changes take place, but that they occur one hundred times as fast as during the spontaneous, non-luminescent oxidation of luciferin in absence of the enzyme. Absorption spectra measured during the luminescent reaction and during the non-luminescent oxidation of luciferin are compared in FIGURE 7.

It seems possible that the shift of absorption maximum from 435 mμ to 465 mμ may be correlated with the reversible oxidation of luciferin demonstrated by Anderson (1936). The two facts that, first, the products of the luminescent reaction of luciferin and of its reversible, non-luminescent oxidation are evidently different (Anderson, 1936) and that, second, the changes in the absorption spectrum are the same during the two reactions (Chase, 1943), would not appear to support such an interpretation. However, the two reactions might be different without necessarily causing a difference in the light-absorbing groups of the molecule. Further experimental evidence is needed before more than tentative conclusions can be drawn. If the assumption be granted that the shift in spectral absorption actually represents the reversible oxidation, then this might indicate the oxidation of a polyhydroxybenzene derivative or related compound. So far, no attempt has been made to demonstrate the reverse shift of the absorption band from 465 mμ to 435 mμ as a result of reduction of spontaneously oxidized luciferin.

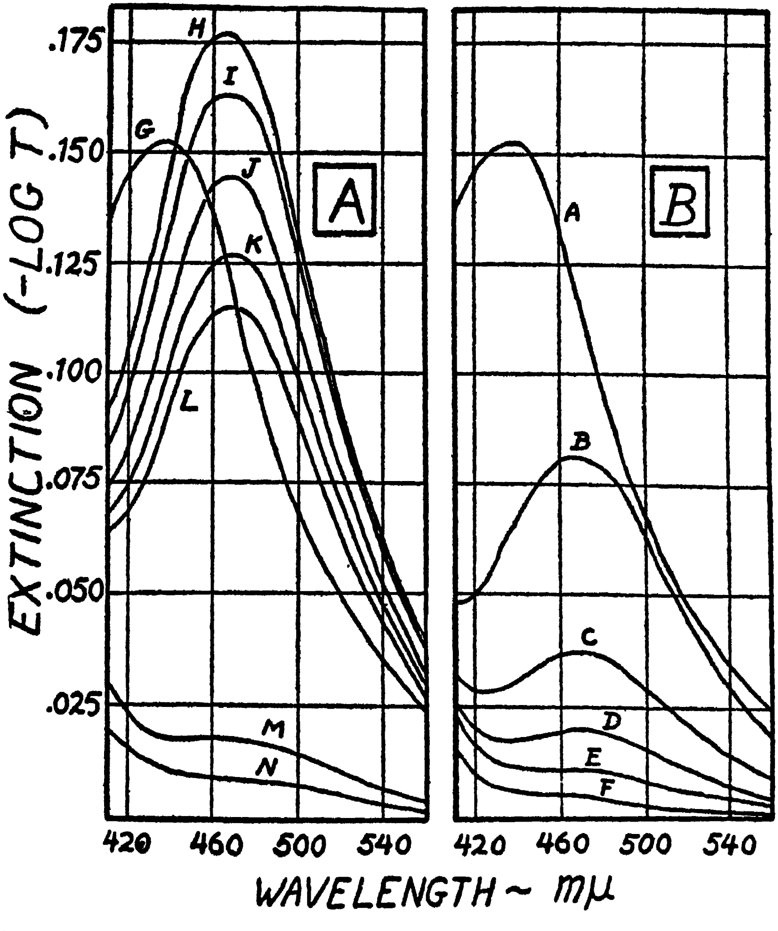


FIGURE 7. The curves shown in the left half of the figure are absorption spectra of a luciferin solution measured during exposure to air. Curve G is the spectrum immediately after the luciferin was dissolved; H, I, J, K, L, M and N, after 9, 20, 30, 39, 49, 455 and 1140 minutes, respectively.

In the right half of the figure are shown absorption spectra of a luciferin solution in which luminescence is occurring due to the action of luciferase, the enzyme. Curve A is the spectrum immediately after the luciferin was dissolved, but before the luciferase was added. Curve B was measured 1 minute after the luciferase was added, curve C after 12 minutes, D after 22 minutes, E after 40 minutes and F after 457 minutes. The changes in the absorption spectrum are about 100 times as rapid during luminescent oxidation by luciferase as during non-luminescent oxidation.

The slow disappearance of the 465 mμ band upon prolonged exposure to dissolved oxygen, or its rapid disappearance in the presence of luciferase, may indicate the disruption of a ring structure. If this be true, it is not surprising that the luminescent oxidation of luciferin should be irreversible. As mentioned earlier, the oxidation products of the naturally occurring polyhydroxybenzene derivatives studied by Ball and Chen (1933) were also unstable compounds. It has been shown by Hooker (1936) that during oxidation by alkaline potassium permanganate, certain 2-hydroxy-1,4-naphthoquinone derivatives undergo an opening of the quinone ring followed by subsequent closure. If a similar reaction occurred in the irreversible oxidation of luciferin (although there is, of course, no experimental evidence for such a mechanism), it seems quite conceivable that the ring, having opened, might become so constituted that it could not close. This would be a possible explanation for the irreversible step in the oxidation, the reversible step having been the oxidation to a quinone.

Chakravorty and Ballentine (1941), in a short note, summarized the then available chemical data on *Cypridina* luciferin and suggested a partial structure for the molecule, illustrated in FIGURE 8. The authors found further evidence for a keto group from the fact that a microcrystalline precipitate was obtained upon treatment of luciferin with hydroxylamine acetate. They suggested that the irreversible luminescent reaction involved the oxidation of the side chain to a carboxyl group, as follows:

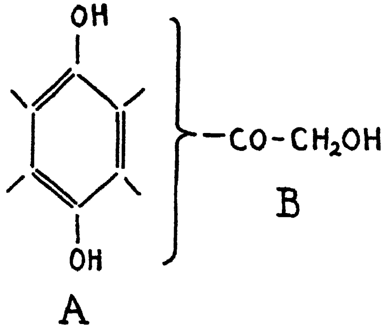


FIGURE 8. The partial structure suggested for luciferin by Chakravorty and Ballentine (1941). The group labeled "A" would represent the part of the molecule capable of reversible oxidation and reduction but not necessarily concerned in the luminescent reaction. The side chain labelled "B" would be capable of reacting with benzoyl chloride or cyanide or, perhaps, with luciferase, and could be oxidized to a carboxyl group as a result of the luminescent reaction with the enzyme.

$$\text{R-CO-CH}_2\text{OH} \xrightarrow[(O_2)]{\text{luciferase}} \text{R-COOH.}$$

If this were correct, they reasoned that it should be possible to regenerate the starting material from the product of the luminescent reaction by appropriate chemical treatment, thus:

$$\text{R-COOH} \xrightarrow{\text{SOCl}_2} \text{R-COCl} \xrightarrow{\text{CH}_2\text{N}_2} \text{R-CO-CHN}_2$$

$$\text{R-CO-CHN}_2 \xrightarrow{\text{dil. H}_2\text{SO}_4} \text{R-CO-CH}_2\text{OH.}$$

After having treated luciferase-oxidized luciferin in this way, they obtained a compound which, they believed, gave luminescence with luciferase, greater in intensity than was obtained from their control. Unfortunately, no quantitative measurement of the luminescence was made, and the light which was obtained was probably only a very small percentage of that which should have been expected, considering the high concentration of luciferase-oxidized luciferin in their starting material. The experiment is especially significant, however, in that it represents probably the first attack upon the problem of the structure of luciferin by the methods of organic chemistry. A continuation of this kind of approach should prove fruitful. Chakravorty and Ballentine also reported microchemical analyses of samples of purified luciferin, which showed no nitrogen, sulfur, halogen, or ash.

McElroy and Ballentine (1944) found evidence for the presence of acid-labile phosphate in purified luciferin. They also reported a significant increase in inorganic phosphate during the luminescent reaction, and stressed the possibility that formation of an energy-rich phosphate bond may be involved in the luminescent reaction. Eymers and van Schouwenburg (1937) had measured the emission spectrum of crude extracts of *Cypridina* luciferin and luciferase. Their analysis showed two fundamental frequencies: one of 21,250, corresponding to a wavelength of 470 mμ, and another smaller component of 18,200, corresponding to a wavelength of 549 mμ. These frequencies represent energies of about 59,000 and 50,000 calories, respectively, and, according to McElroy's and Ballentine's calculations, the preservation of energy as phosphate bond energy from oxidation of the postulated side chain of luciferin, and the subsequent release of this energy through reaction with luciferase would be sufficient to produce luminescence having the observed spectral quality.

If luciferin does have an hydroxybenzene, or related, structure and if it can be reversibly oxidized to a quinone, measurements of its ultraviolet absorption spectrum would be expected to throw light on the specific configuration of the molecule. The only such measurements so far recorded (Chase, 1943; Chase and Giese, 1940) cannot be regarded too seriously, since impurities were certainly present in the luciferin solutions even after two cycles of purification by Anderson's procedure (1935). Impurities would doubtless cause much greater interference with the ultraviolet spectrum than in the visible. An ultraviolet and visible absorption spectrum of luciferin in pH 6.8 phosphate buffer is shown in FIGURE 9, measured immediately after dissolving the luciferin and after exposure to air during two time-intervals. A qualitative resemblance to the spectra of *p*-quinone and of certain naphtho- and anthraquinone derivatives actually seems to exist, and it is possible that luciferin may have an aromatic ring structure as its nucleus. Since the solvent frequently exerts a considerable effect on the absorption spectrum of a solution, precise measurements of the absorption spectrum of luciferin should be made, using a variety of solvents, in order to secure data susceptible of interpretation in terms of structure.

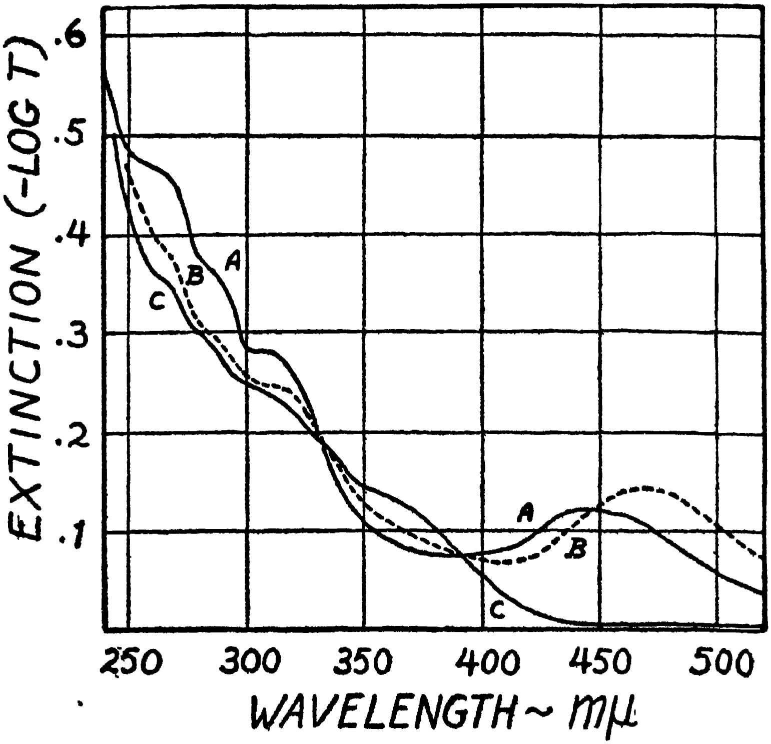


FIGURE 9. Visible and ultraviolet absorption spectra of a luciferin solution at pH 6.8. The spectra were recorded with the Harrison and Bentley spectrophotometer (1940). Curve A was measured as soon as possible after the luciferin had been dissolved. Curve B was measured after exposure of the solution to air for 32 minutes, and curve C after 24 hours.

With the ultraviolet absorption spectrum measurements, such as they are, favoring *p*-quinone, naphthoquinone or anthraquinone derivatives, the first-mentioned structure seems most acceptable in view of the fact that the E_0' (pH 7) of the luciferin system is almost identical with that of quinhydrone (Anderson, 1936; Korr, 1936) and not sufficiently negative to fall in the range characteristic of naphtho- and anthraquinones. It must be remembered, however, that substitution in a quinone ring can greatly change the oxidation-reduction potential in either a positive or negative direction, depending upon the nature of the substituent group.

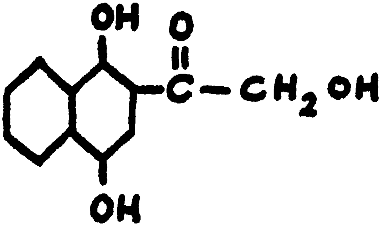
As stated earlier, luciferin and luciferase have not yet been extracted from luminous bacteria although there is every reason to believe that they occur. Recently Kluyver, van der Kerk, and van der Burg (1942) determined the action spectrum for the inhibition of bacterial luminescence by light of various wavelengths in the visible and ultraviolet. Assuming that this action spectrum represented the absorption spectrum of a compound closely related to bacterial luciferin, probably oxidized luciferin, they tentatively identified this compound from its spectrum, and from other considerations, as a 1,4-naphthoquinone derivative with a ketohydroxy group directly substituted in the quinone ring. They based their interpretation largely upon the conclusions of Chakravorty and Ballentine (1941), discussed above. There is a possibility, as the authors point out, that the action spectrum in this case may be affected by inert colored compounds in the medium or in the bacterial cell, or that it may represent the absorption spectrum of some compound which is so acting as to sensitize to light a component of the luminescent reaction. The action spectrum as measured might not, therefore, necessarily represent the absorption spectrum of oxidized luciferin. Although the actual comparison of the measured action spectrum with the absorption spectrum of the naphthoquinone derivative is not shown in their paper, it does appear in the doctorate thesis of van der Kerk (1942). While there is a qualitative resemblance between the two spectra as far as shape is concerned, the actual positions of absorption bands differ rather widely. Furthermore, the spectrum of the naphthoquinone derivative was measured in hexane, whereas the solvent for the hypothetical oxidized bacterial luciferin is unknown but is presumably largely aqueous. For these reasons, the conclusions as to structure should, as the authors themselves stress, be considered as purely tentative.

Another investigation on the luminescent reaction in bacteria which has yielded information as to certain characteristics of bacterial luciferin is that of Johnson, van Schouwenburg, and van der Burg (1939). They recorded the flash of luminescence in bacteria, following anaerobiosis under a variety of conditions. They concluded* that, since their data indicated that the luminescent oxidation of luciferin by luciferase and oxygen is evidently irreversible, whereas the dark oxidation by a substrate (or its breakdown product) is reversible, the situation is rather closely comparable to that in *Cypridina* extracts. Anderson (1936) had shown that the luminescent and non-luminescent oxidations of *Cypridina* luciferin resulted in different products. Johnson, van Schouwenburg, and van der Burg suggested, in addition, that the two oxidations might well occur on different portions of the luciferin molecule. That is, that the molecule might have one group which might easily undergo reversible oxidation and reduction, while an entirely different group might be irreversibly oxidized with light emission.

---

* See page 211 of their paper.

A privately printed doctoral thesis by C. J. P. Spruit (1946) has recently appeared. This contains, among other things, the ultraviolet absorption spectra of a number of compounds with structures related to those that have been suggested for luciferin as a result of the work of Anderson (1936), Giese and Chase (1940), Chakravorty, and Ballentine (1941), van der Kerk (1942), Kluyver, van der Kerk, and van der Burg (1942), and Chase (1943). Spruit concludes that bacterial luciferin may have the following structure:



The compounds whose absorption spectra he measured include some thirty-three anthra-, naphtho- and benzoquinone and hydroquinone derivatives. Since Spruit's conclusion as to the structure of bacterial luciferin depends upon a comparison of the spectra of these known compounds with that of luciferin, the absorption spectrum of luciferin itself should be established more definitely than has so far been possible before the structure suggested by Spruit can be considered proven. The ultraviolet absorption spectra so far reported for *Cypridina* and bacterial luciferin, whether measured by direct or indirect methods, may easily have been distorted by the presence of impurities. Consequently, comparisons with the spectra of known compounds must be made with caution and a good deal of reservation.

Johnson (1947) in a comprehensive review paper on bacterial luminescence and Johnson and Eyring (1944) have called attention to the possibility that luciferin may be a prosthetic group on the protein enzyme, luciferase. Part of their evidence for this hypothesis is their demonstration that *Cypridina* luciferase solutions which have been subjected to prolonged dialysis can be made to luminesce by the addition of $Na_2S_2O_4$, followed by aeration of the solutions. They suggest that this prosthetic group may possibly be of flavin nature, although most of the experimental evidence available at present would appear not to favor such a structure for the luciferin molecule (Anderson and Chase, 1944; McElroy and Ballentine, 1944).

The purification of *Cypridina* luciferin by crystallization has been a goal of all workers in the chemical field of bioluminescence since luciferin was first extracted but, save for one doubtful preliminary report, no one has succeeded in obtaining crystals. Kanda (1932) reported crystallization of luciferin but his paper did not contain photographs of the crystals. The work has not been confirmed by other investigators nor have further papers on the subject been published by him.

The only crystalline derivative of luciferin so far reported is the microcrystalline precipitate mentioned by Chakravorty and Ballentine (1941), which was obtained on treatment with hydroxylamine acetate. Here, too, no photographs are shown. When the irreversible reaction between purified luciferin and cyanide (Giese and Chase, 1940) was first observed, indicating the possible presence in luciferin of an aldehyde or keto group, an attempt was made to obtain a crystalline derivative by reaction with 2,4-dinitrophenylhydrazine. However, this was not successful. Obviously, the formation of any kind of crystalline derivative, whether capable of producing light with luciferase or not, would be valuable as a means of obtaining material of a high degree of purity for quantitative analysis, as a lead towards establishment of structure. The scarcity of the starting material, *Cypridina,* means that micro-methods, rather than the easier macro-techniques, must be used in all such work.

It may be interesting to summarize briefly some unpublished data by Dr. Harvey and myself on relative solubilities of *Cypridina* luciferin, both crude and purified, in various solvents. The amount of luciferin in solution was determined by measuring the total light available from an aliquot of the solution, on addition of luciferase, at constant pH. As had been observed earlier by Harvey, by Kanda, and by Anderson, we, too, found that great differences in the relative solubility in different solvents existed, depending upon whether the crude, dry *Cypridina* powder or the purified luciferin were used. The purified luciferin (Anderson's method), in the quantity used, was found to be completely soluble in the following solvents: methanol, ethanol, butyl alcohol, isobutyl alcohol, and propyl alcohol. Amyl alcohol, acetone, chloroform, aniline, and 0.06 *M,* pH 6 8 phosphate buffer dissolved only 50-75 per cent of the available luciferin. Ethyl ether, petroleum ether, and benzene dissolved none.

Luciferin in the form of the dry, powdered *Cypridina* material was found to be most soluble in methanol. Calling this value 100 per cent, ethanol dissolved about 16 per cent, propyl alcohol about 10 per cent, and butyl alcohol, isobutyl alcohol, amyl alcohol, acetone, chloroform, and ethyl ether dissolved practically none (0.5 per cent). Preliminary benzene extraction of the crude, powdered material in a Sohxlet caused a significant increase in the relative amounts of luciferin extractable by ethanol, butyl alcohol, and propyl alcohol. No measurable luciferin was removed by the benzene in the preliminary treatment. The details are given in TABLE 1.

TABLE 1

RELATIVE SOLUBILITIES IN VARIOUS SOLVENTS OF PURIFIED LUCIFERIN, LUCIFERIN IN DRY *Cypridina* POWDER, AND LUCIFERIN IN BENZENE-EXTRACTED DRY *Cypridina* POWDER

Solvent	Relative amount of luciferin extracted		
	Purified luciferin	Untreated *Cypridina* powder	Benzene-extracted powder
Methyl alcohol	98	100	100
Ethyl alcohol	94	16	34
Butyl alcohol	100	4	12
Isobutyl alcohol	100	5	—
Propyl alcohol	98	10	24
Amyl alcohol	83	0	5
Acetone	74	1	0
Chloroform	44	3*	6*
Ethyl ether	0	0	2
Petroleum ether	0	—	—
Benzene	0	—	—
Aniline	77	—	—
Phos. buff., pH 6.8	49	—	—

* Solid fragments present in the solution.

It seems quite clear, from these results on relative solubilities, that the luciferin, as it occurs in the crude state in the gland of *Cypridina,* must be bound in some way, or else be actually different, chemically, than it is in the purified condition. This same conclusion was reached earlier by both Harvey and Kanda. If there is an actual chemical difference, this cannot be great enough to alter the structure sufficiently to preclude reaction with the enzyme, luciferase. These differences between the unpurified and purified material certainly make it evident that experimental work must be confined to only one kind of luciferin preparation: preferably, at least at this stage, to luciferin purified by Anderson's method. Otherwise, inconsistencies are bound to occur which will confuse, rather than help to clarify, the problem.

As will have become apparent, the chemical structure of the luciferin of *Cypridina* is still far from known, although a number of facts have accumulated that bear directly on the problem, and several partial structures have been postulated on the bases of these data. It seems likely that the reversible oxidation may represent a reaction analogous to the oxidation of hydroquinone. Furthermore, the redox potential of the luciferin system certainly resembles those of naturally occurring polyhydroxybenzene derivatives whose oxidized forms are unstable, as is also the case with luciferin. The irreversible reaction between purified luciferin and cyanide may indicate a free aldehyde or keto group, and the chemical experiments of Chakravorty and Ballentine point to some sort of ketohydroxy side chain. The reaction with azide may indicate a quinone nucleus for reversibly oxidized luciferin, although this interpretation is only one of several which could be made in the light of those data. The evidence for a naphthohydroquinone structure with a ketohydroxy side chain for bacterial luciferin is probably not sufficiently conclusive to do more than indicate such a possibility. The published measurements of the ultraviolet absorption spectrum of *Cypridina* luciferin, although they, too, suggest an aromatic ring structure as the skeleton of the molecule, suffer from impurities and require remeasurement under a variety of conditions. The splitting-off of phosphate in the luminescent reaction certainly seems relevant in a consideration of structure.

A promising lead for determining the structure of luciferin may be its rather peculiar color change during oxidation. The change in spectral absorption from 435 mμ to 465 mμ on exposure to air and the subsequent loss of all visible light absorption on further exposure are rather unique properties. Any known compound which possesses such properties should be capable of transformation, by suitable chemical treatment, into luciferin. Since the luminescent reaction itself is enzyme-catalyzed, it is obvious that, in addition to the fundamental structure of the molecule, there must be one or more very specific groupings. The problem, then, once the basic structure has been found from absorption spectrum data, would involve making various derivatives by substitution of side chains, until a compound is obtained which can react with luciferase to give luminescence.

BIBLIOGRAPHY

Amberson, W. R.

1922. Kinetics of the bioluminescent reaction in *Cypridina*. I and II. J. Gen. Physiol. **4**: 517, 535.

Anderson, R. S.

1933. The chemistry of bioluminescence. I. Quantitative determination of luciferin. J. Cell. & Comp. Physiol. **3**: 45.
1935. Studies on bioluminescence. II. The partial purification of *Cypridina* luciferin. J. Gen. Physiol. **19**: 301.
1936. Chemical studies on bioluminescence. III. The reversible reaction of *Cypridina* luciferin with oxidizing agents and its relation to the luminescent reaction. J. Cell. & Comp. Physiol. **8**: 261.

Anderson, R. S., & A. M. Chase

1944. The nature of *Cypridina* luciferin. J. Am. Chem. Soc. **66**: 2129.

Ball, E. G., & T. T. Chen

1933. Studies on oxidation-reduction potentials. XX. Epinephrine and related compounds. J. Biol. Chem. **102**: 691.

Chakravorty, P. N., & R. Ballentine

1941. On the luminescent oxidation of luciferin. J. Am. Chem. Soc. **63**: 2030.

Chance, B., E. N. Harvey, F. Johnson, & G. Millikan

1940. The kinetics of bioluminescent flashes. A study in consecutive reactions. J. Cell. & Comp. Physiol. **15**: 195.

Chase, A. M.

1940. Changes in the absorption spectrum of *Cypridina* luciferin solutions during oxidation. J. Cell. & Comp. Physiol. **15:** 159.

1942. The reaction of *Cypridina* luciferin with azide. J. Cell. & Comp. Physiol. **19:** 173.

1943. The absorption spectrum of luciferin and oxidized luciferin. J. Biol. Chem. **150:** 433.

1945. The visible absorption band of reduced luciferin. J. Biol. Chem. **159:** 1.

Chase, A. M., & A. C. Giese

1940. Effects of ultraviolet radiation on *Cypridina* luciferin and luciferase. J. Cell. & Comp. Physiol. **16:** 323.

Chase, A. M., & P. B. Lorenz

1945. Kinetics of the luminescent and non-luminescent reactions of *Cypridina* luciferin at different temperatures. J. Cell. & Comp. Physiol. **25:** 53.

Eymers, J. G., & K. L. van Schouwenburg

1937. On the luminescence of bacteria. III. Further quantitative data regarding spectra connected with bioluminescence. Enzymologia **3:** 235.

Fieser, L. F., & J. L. Hartwell

1935. The reaction of hydrazoic acid with the naphthoquinones. J. Am. Chem. Soc. **57:** 1482.

Fisher, K. C., & R. Oehnell

1940. The steady state frequency of the embryonic fish heart at different concentrations of cyanide. J. Cell. & Comp. Physiol. **16:** 1.

Giese, A. C., & A. M. Chase

1940. The effect of cyanide on *Cypridina* luciferin. J. Cell. & Comp. Physiol. **16:** 237.

Hardy, A. C.

1935. A new recording spectrophotometer. J. Opt. Soc. Am. **25:** 305.

Harrison, G. R., & E. P. Bentley

1940. An improved high speed recording spectrophotometer. J. Opt. Soc. Am. **30:** 290.

Harvey, E. N.

1917. Studies on bioluminescence. IV. The chemistry of light production in a Japanese ostracod crustacean, *Cypridina hilgendorfii* Müller. Am. J. Physiol. **42:** 318.

1925. The inhibition of *Cypridina* luminescence by light. J. Gen. Physiol. **7:** 679.

1926. Further studies on the inhibition of *Cypridina* luminescence by light, with some observations on methylene blue. J. Gen. Physiol. **10:** 103.

1940. Living Light. Princeton University Press. Princeton, N. J.

1941. Review of bioluminescence. Ann. Rev. Biochem. **10:** 531.

1948. Introductory remarks: A general survey of bioluminescence. Ann. N. Y. Acad. Sci. **49** (3): 329.

Hooker, S. C.

1936. On the oxidation of 2-hydroxy-1,4-naphthoquinone derivatives with alkaline potassium permanganate. J. Am. Chem. Soc. **58:** 1174.

Johnson, F. H.

1947. Bacterial Luminescence. In: *Advances in Enzymology* **7:** 215. Interscience Publishers, Inc. New York.

Johnson, F. H., & H. Eyring

1944. The nature of the luciferin-luciferase system. J. Am. Chem. Soc. **66:** 848.

Johnson, F. H., H. Eyring, & R. W. Williams

1942. The nature of enzyme inhibitions in bacterial luminescence: sulfanilamide, urethane, temperature and pressure. J. Cell. & Comp. Physiol. **20:** 247.

Johnson, F. H., K. L. van Schouwenburg, & A. van der Burg
1939. The flash of luminescence following anaerobiosis of luminous bacteria. Enzymologia **7:** 195.

Kanda, S.
1924. Physico-chemical studies on bioluminescence. V. The physical and chemical nature of the luciferine of *Cypridina hilgendorfii.* Am. J. Physiol. **68:** 435.
1929. Physico-chemical studies on bioluminescence. VII. The solubility of *Cypridina* luciferin in organic solvents. Sci. Papers Inst. Phys. & Chem. Research, Tokyo **10:** 91.
1932. Crystalline luciferin. Suppl. Sci. Papers Inst. Phys. & Chem. Research, Tokyo **18:** 1.

Kluyver, A. J., G. L. M. van der Kerk, & A. van der Burg
1942. The effect of radiation on light emission by luminous bacteria. I and II. Proc. Nederl. Akad. Wetenschappen **45:** 886, 962.

Korr, I. M.
1936. The luciferin-oxyluciferin system. J. Am. Chem. Soc. **58:** 1060.

McElroy, W. D., & R. Ballentine
1944. The mechanism of bioluminescence. Proc. Nat. Acad. Sci. **30:** 377.

Spruit, C. J. P.
1946. Naphthochinonen en Bioluminescentie. Doctorate thesis. Drukkerij Fa. Schotanus & Jens. Utrecht.

van der Kerk, G. J. M.
1942. Onderzoekingen over de Bioluminescentie der Lichtbacteriën. Doctorate Thesis. N. V. Kemink en Zoon. Utrecht.

THE FUNDAMENTAL ACTION OF PRESSURE, TEMPERATURE, AND DRUGS ON ENZYMES, AS REVEALED BY BACTERIAL LUMINESCENCE

By Frank H. Johnson and Henry Eyring
Department of Biology, Princeton University, Princeton, N. J., and
The Graduate School, University of Utah, Salt Lake City, Utah

With any given enzyme reaction or more complex biological process, a full interpretation of the kinetics involves an understanding of the mechanisms through which temperature, hydrostatic pressure, and inhibitors of various kinds influence the observed rate. While the theory of equilibria in regard to temperature, pressure, and concentration of reactants has been known for some years, it is scarcely a decade since the rational and precise theory for change became available, in the theory of absolute reaction rates[1, 2] for a rate process such as that of a simple chemical reaction. Inasmuch as the same theory undoubtedly holds as well for each of the individual reactions which collectively lead to any biological phenomenon, the problem is to see how it may be applied to single reactions in the midst of many, and to what extent it may account, quantitatively, for the process as a whole. In some cases, for example with an extracted and purified enzyme system, the kinetics might be expected to be relatively uncomplicated, although simultaneous reactions involving the same molecule may influence the over-all rate of the measured process. In living cells, on the other hand, a series of systems is generally concerned, and under various conditions one or more systems may be largely rate-determining, in each case influenced in their activity by simultaneous equilibria and rate processes. When a single system remains largely limiting, however, it is reasonable to believe that even complex processes might be analyzed, approximately, in the manner of a single reaction.

For investigating the problem at hand, luminescence possesses distinct, and in certain respects unique advantages. In the first place, the intensity of the light is evidently proportional to the reaction velocity of the "luciferase" with "luciferin," as shown by kinetic studies with crude as well as with partially purified extracts of *Cypridina*.[3] Although luminescent extracts have not as yet been obtained from bacteria, spectroscopic and other lines of evidence indicate that the light-emitting system is not fundamentally different in the two cases. Inasmuch as the intensity of luminescence may be easily recorded by means of a photoelectric cell or other devices,[4] the rate of reaction, in relative units, can be determined with considerable accuracy for a given instant. Likewise, the course of more or less rapidly changing reaction rates may be followed during very brief as well as over longer periods of time.

In non-reproducing bacteria, under favorable physiological conditions, luminescence is in a steady state, with constant intensity. In extracts of *Cypridina,* the reaction is first order with respect to both the concentration of dihydroluciferin and of active luciferase, and the luminescent oxidation is accompanied by the destruction of a large part of the luciferin.[3, 5] In bacteria, if any considerable destruction of the luciferin accompanies the light-emitting oxidation, the luciferin must be formed from some precursor at the same rate, in order for a uniform intensity to be maintained. Since excited molecules are generally stabilized by radiating, whereas excited molecules which do not radiate are apt to be destroyed, repeated oxidation and reduction of the same luciferin molecules very likely occurs, with perhaps much less of the destructive reaction in living cells than in extracts. In any case, the uniform luminescence intensity of bacteria indicates that the concentration of the reactants, luciferin and luciferase, in effect remains constant with time. Thus, the intensity (I) of the light will be proportional to the amount of active luciferase (A_n) times the amount of reduced luciferin (LH_2) times the specific reaction rate constant, k, times some proportionality constant, b:

$$I = bk(A_n)(LH_2). \qquad (1)$$

While it is apparent that several equilibria and rate processes precede actual light emission[6] the evidence from flow-method studies with *Cypridina* extracts[7] indicates that EQUATION 1 represents the slowest reaction. A scheme of consecutive reactions consistent with the known facts concerning luminescence in bacteria and in extracts is given in TABLE 1. Excited luciferin is designated by L^*, and destroyed luciferin by L_1.

TABLE 1

	Reactions with luciferase		Additional reactions that occur with and without luciferase
(1)	$AL + XH_2 \rightleftharpoons ALH_2 + X$	(1′)	$L + XH_2 \rightleftharpoons LH_2 + X$
(2)	$A + LH_2 \rightleftharpoons ALH_2$		
(3)	$ALH_2 + O_2 \rightleftharpoons ALH + HO_2$	(3′)	$LH_2 + O_2 \rightleftharpoons LH + HO_2$
(4)	$ALH \rightleftharpoons AL^- + H^+$	(4′)	$LH \rightleftharpoons L^- + H^+$
(5) alpha	$AL^- + LH \rightarrow AL^* + LH^- \rightarrow AL + LH^- + h\nu$	(5′) alpha	$L^- + LH \rightarrow L^* + LH^- \rightarrow L + LH^- + h\nu$
(5) beta	$AL^- + LH \rightarrow AL + LH^-$	(5′) beta	$L^- + LH \rightarrow L + LH^-$
(5) gamma	$AL^- + LH \rightarrow AL_1 + LH^-$	(5′) gamma	$L^- + LH \rightarrow L_1$
(6)	$ALH + O_2 \rightarrow AL + HO_2$	(6′)	$LH + O_2 \rightarrow L + HO_2$
(7)	$AL + O_2 \rightarrow AL_1$	(7′)	$L + O_2 \rightarrow L_1$

In this scheme, the transfer of an electron between two semi-quinone forms of the luciferin leads to excitation, followed by radiation. The luciferin is presumed to be the radiating molecule, because of the correspondence between the absorption spectrum of luciferin and the emission spectrum of the luminescent reaction,[8, 9] together with the fact that luciferin emits light in 95 per cent alcohol at 70° C.,[10] a condition under which the enzyme might be expected to be inactive.

Thus, the over-all process of luminescence may be limited in various ways, even though reaction 2 ordinarily remains the slowest of the series. At very low oxygen tensions, for example, reaction 3 becomes limiting. In acid pH, the amount of ALH decreases and, to this extent, reaction 4 may be considered limiting. In alkaline pH, the amount of *LH* may be considered limiting, etc. All these reactions, however, are fast in comparison with reaction 2, even though alterations in the steady state concentration of subsequent reactants, *e.g.*, by changes in pH, influence the level of luminescence intensity in a corresponding manner. At optimum pH, with oxygen not limiting, and with excess of glucose, the over-all rate is determined by the specific reaction rate constant, the amount of luciferin, and the amount of active luciferase, as given in EQUATION 1.

A second, unusual advantage of luminescence in analyzing the kinetics of over-all processes that occur in living cells is that the enzyme system concerned is not preceded by a considerable number of other systems engaged in a stepwise hydrogen or electron transfer. The evidence that this is so derives from the facts that (a) glucose added to washed cells immediately results in large increases in luminescence, showing that some of the hydrogen from glucose is transferred *via* the luciferin-luciferase system, with light emission;[11, 12, 13] and (b) the wavelength of maximum intensity corresponds to a transition with $\Delta F^\circ = 60{,}700$ calories, as compared with the average ΔF° of 57,340 calories for two hydrogens in the oxidation of glucose.[6] Thus, in luminescence, some fraction of the hydrogens from glucose goes almost directly to oxygen, by way of the luminescent system, and the energy is released through visible radiation.

Finally, since luminescence may be studied both in living bacterial cells and in a corresponding system extracted and partially purified from *Cypridina,* it is possible to ascertain whether inhibitors which influence the intensity of the light in cells also act, at equivalent concentrations, on the system *in vitro*. Moreover, by measuring the effect of a given inhibitor on the total light in extracts, it is possible to distinguish between those which, like certain ions[14] quench luminescence, and those, such as urethane or sulfanilamide,[15] which retard the rate of the enzyme reaction without affecting the total light produced.

In order to account for the reversible effects of temperature and hydrostatic pressure on bacterial luminescence, it is necessary to modify EQUATION 1 to include an equilibrium between active and inactive forms of one of the reactants.[16, 17] Since this equilibrium, which rapidly becomes a conspicuous limiting factor with rise in temperature beyond the normal optimum, has been found to be characterized by the high heat and entropy typical of protein denaturation, it evidently concerns the enzyme.[18, 19, 20] Letting K_1 represent the equilibrium constant between active (A_n), and inactive or reversibly denatured (A_d) forms of the enzyme, and A_0 the total of $(A_n) + (A_d)$, EQUATION 1 becomes

$$I_1 = \frac{bk(LH_2)(A_0)}{1 + K_1}, \tag{2}$$

which, by definition of the equilibrium constant K_1 and of the rate constant k in accordance with the theory of Absolute Reaction Rates[1, 2] may be written:

$$I_1 = \frac{b\kappa \frac{kT}{h} e^{-\frac{\Delta F\ddagger}{RT}} (LH_2)(A_0)}{1 + e^{-\frac{\Delta F_1}{RT}}}$$

$$= \frac{b\kappa \frac{kT}{h} e^{-\frac{\Delta H\ddagger}{RT}} e^{\frac{\Delta S\ddagger}{R}} (LH_2)(A_0)}{1 + e^{-\frac{\Delta H_1}{RT}} e^{\frac{\Delta S_1}{R}}}$$

$$= \frac{b\kappa \frac{kT}{h} e^{-\frac{\Delta E\ddagger}{RT}} e^{-p\frac{\Delta V\ddagger}{RT}} e^{\frac{\Delta S\ddagger}{R}} (LH_2)(A_0)}{1 + e^{-\frac{\Delta E_1}{RT}} e^{-p\frac{\Delta V_1}{RT}} e^{\frac{\Delta S_1}{R}}}. \tag{3}$$

EQUATION 3 contains some unknown quantities, (LH_2), (A_0), and $\Delta S\ddagger$, which cannot be readily determined. Under chosen conditions, however, where luminescence of fully aerated, non-reproducing cells is constant with time, these quantities may be assumed essentially constant. The same applies to κ, the transmission coefficient, k, the Boltzman constant, h, Planck's constant, and b, the proportionality constant. Consequently, they may be included in a single constant, c, and EQUATION 3 then becomes

$$I_1 = \frac{cTe^{-\frac{\Delta H\ddagger}{RT}}}{1 + e^{-\frac{\Delta H_1}{RT}} e^{\frac{\Delta S_1}{R}}} = \frac{c'Te^{-\frac{\Delta E\ddagger}{RT}} e^{-p\frac{\Delta V\ddagger}{RT}}}{1 + e^{-\frac{\Delta E_1}{RT}} e^{-p\frac{\Delta V_1}{RT}} e^{\frac{\Delta S_1}{R}}} \tag{4}$$

EQUATIONS 3 and 4 show that the over-all intensity of luminescence, under optimal physiological conditions of pH, oxygen, salt concentration, etc., will be determined by the influence of temperature on two reactions involving the same molecule, namely, the rate of the catalytic reaction and the inactivation equilibrium of the enzyme, respectively.

The values of the constants in EQUATION 4 may be estimated from experiments in which I is measured at different temperatures and pressures. Thus, the apparent $\Delta H^{\ddagger}$, which at normal pressure is indistinguishable from "μ" of the Arrhenius equation, is computed from the slope of the line relating $ln\ I$ and $\frac{1}{T}$ in the low temperature range where the value of K_1 is negligible. ΔH_1 is obtained from the slope of the decreasing I with increasing T at temperatures above the optimum, and adding to this the value of $\Delta H^{\ddagger}$, disregarding signs. $\Delta V^{\ddagger}$ may be determined from the slope of the line relating $ln\ I$ to pressure (p) in atmospheres at low temperatures. ΔV_1 is obtained from a similar plot at high temperatures, ΔE from the relation, $\Delta H = \Delta E + p\Delta V$, and ΔS from the relation

$$K = e^{-\frac{\Delta H}{RT}} e^{\frac{\Delta S}{R}} = e^{-\frac{\Delta E}{RT}} e^{-p\frac{\Delta V}{RT}} e^{\frac{\Delta S}{R}} .$$

These constants then make it possible to calculate, using EQUATION 4, the intensity of luminescence with some accuracy at various temperatures and pressures, as shown (FIGURE 1) by Eyring and Magee.[19] In this

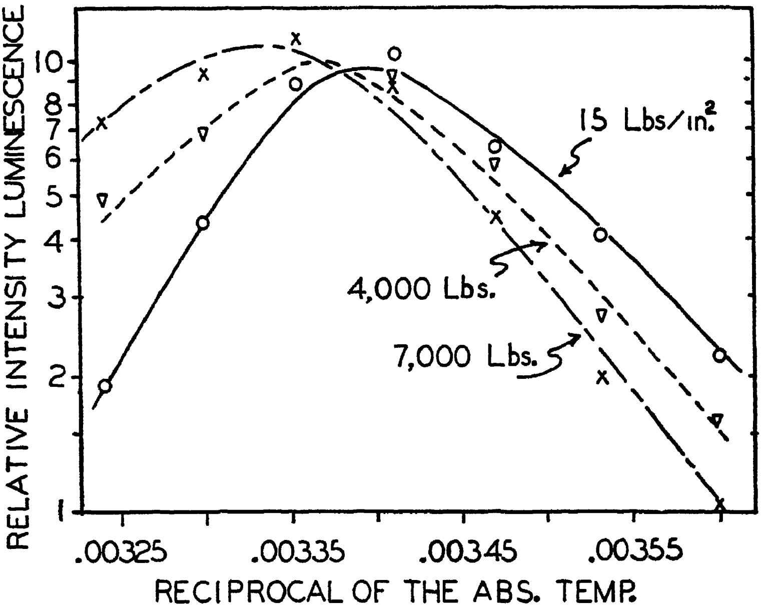


FIGURE 1. Relation between the intensity of luminescence (*Photobacterium phosphoreum*) and temperature, under normal and increased hydrostatic pressures of 4,000 and 7,000 lbs./sq. in., respectively. The solid lines are curves calculated by Eyring and Magee[19] as described in the text; the points represent data from experiments of Brown, Johnson, and Marsland.[17] Logarithmic scale on the ordinate.

case, $\Delta V^\ddagger$ amounts to 50 cc. per mole at 0° C, and ΔV_1 to 64 cc. at 35° C, but a temperature dependence of these constants has been taken into account. The constants employed are as follows: $\Delta E^\ddagger = 17{,}220$ calories, $\Delta V^\ddagger = 546.4 - 1.813T$ cc., $\Delta E_1 = 55{,}260$ calories, $\Delta S_1 = 184$ Entropy units, $\Delta V_1 = -922.8 + 3.206T$ cc.

The difference in the values of the constants for the luciferin-luciferase reaction and for the inactivation equilibrium of luciferase, respectively, are largely responsible for the optimal temperature of luminescence at a given pressure, and likewise for an optimal pressure at a given temperature. The apparent activation energy of 17,220 calories at atmospheric pressure is similar to, or not very different from, those familiarly associated with enzyme reactions and more complex biological processes as a whole. The heat and entropy of the inactivation equilibrium, as pointed out above, are typical of protein denaturation. The volume changes of activation and of reaction, respectively, are of particular interest, since both are high and show that in each case the reaction involves a very large molecule. These volume changes indicate that fairly drastic alterations in the structure of the molecule take place both in the formation of the activated complex and in the reversible inactivation. They probably indicate considerable unfolding of the protein from a somewhat globular to a more fibrous form. Thus, in the process of catalysis, it would appear necessary for the enzyme to change its configuration to fit the substrate, although such a change would not be necessary if the active or combining groups of the enzyme are at the surface, which appears to be the situation with invertase.[21]

Turning now to the action of inhibitors which combine reversibly with the luminescent system, it is evident from EQUATION 2 that the equilibrium through which the system is inactivated may be established either independently of the denaturation equilibrium (Type I) or in relation to it (Type II). The former mechanism has the likeness of a combination of the inhibitor with a prosthetic, group of the enzyme, or possibly the luciferin. In this case, kinetic data alone are not sufficient to distinguish between the enzyme and substrate as site of action.[18] Type II represents a combination of the inhibitor with bonds which are involved in the denaturation equilibrium. Although Types I and II cannot be distinguished experimentally on the basis of the relation between concentration of inhibitor and the amount of inhibition observed, other conditions remaining constant, they may, in general, be expected to behave in different ways as pressure and temperature are varied. Thus, in Type I, letting K_2 represent the equilibrium constant, independent of K_1, pressure would be expected to have only slight influence, while a rise in temperature should result in a decrease in per cent inhibition, as the enzyme-inhibitor complex is dissociated and K_2 becomes smaller. Since the inhibition at a given drug concentration is less at lower temperatures, an increase in apparent activation energy of the luminescent reaction occurs in the presence of the drug, and the temperature of maximum luminescence is slightly increased. On the other hand, in Type II, with equilibrium constant K_3 because of its dependence on K_1, which is characterized by a large volume change of reaction, a pronounced effect of pressure on the inhibition would be anticipated. Thus, at temperatures near the normal optimum, moderate inhibitions by alcohol can be completely reversed by hydrostatic pressure,[6] as illustrated in FIGURE 2. The influence of pressure on the inhibition caused by various drugs, including representatives of both types, is illustrated in FIGURE 3.[22] As for temperature, the influence in this type depends on both K_1 and K_3, that is: although the enzyme-inhibitor complex dissociates with rise in temperature, the bonds with which the inhibitor combines may be more available; K_1 increases and K_3 decreases as the temperature is raised, and the net result depends on K_1 times K_3. In the cases studied, the per cent inhibition increases with rise in temperature, resulting in a lowering of the apparent activation energy, a decrease in the observed heat of denaturation, and a

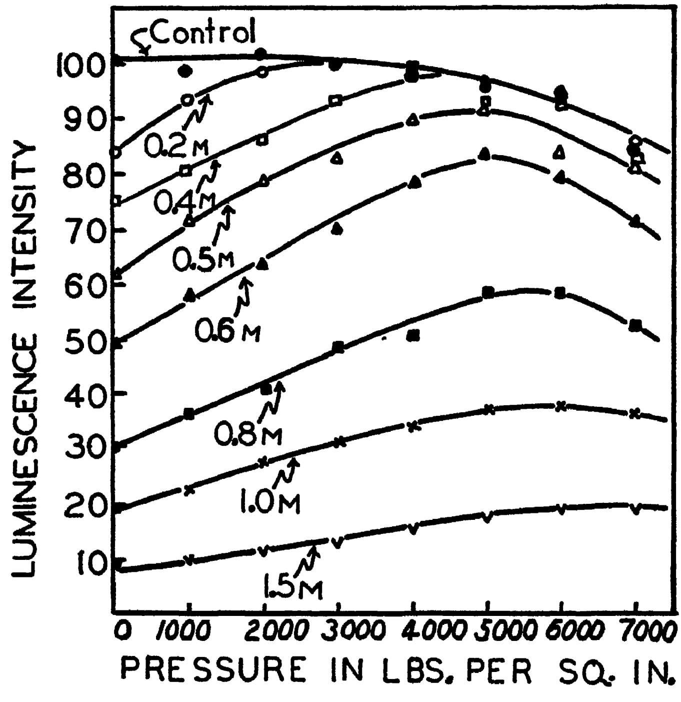


FIGURE 2. The relation between hydrostatic pressure and the amount of inhibition of luminescence at 17.5° C caused by the various concentrations of alcohol indicated on the figure.[6] (*Photobacterium phosphoreum.*)

somewhat lower temperature of maximum luminescence, in the presence of a given drug concentration.

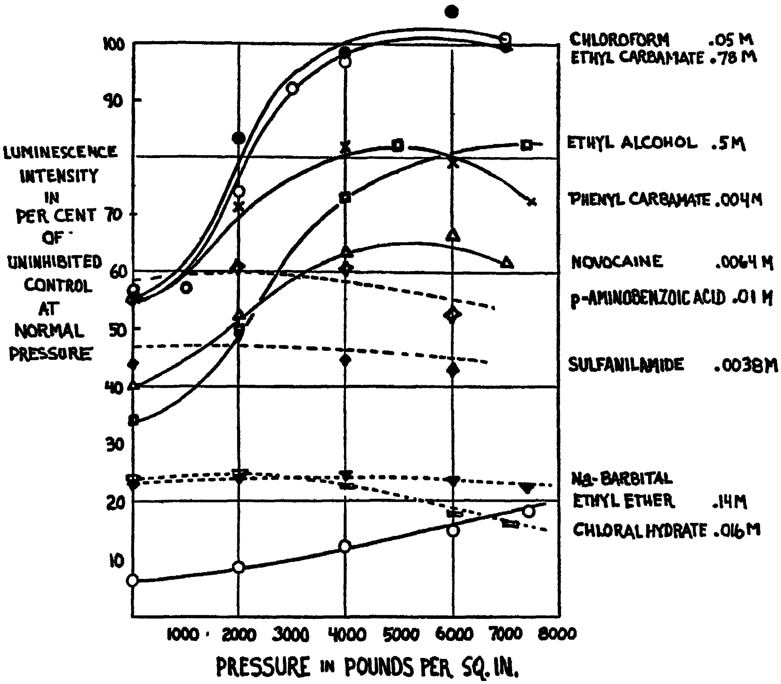


FIGURE 8. The relation between hydrostatic pressure and the amount of inhibition of luminescence caused by a given concentration of various drugs, at 17 to 18° C. The intensity of the drug-free control at atmospheric pressure has arbitrarily been taken as 100, and the other points computed in relation to it. (*Photobacterium phosphoreum.*[22])

The relation between luminescence intensity and temperature, without the addition of drugs, and in the presence of a single concentration of sulfanilamide and urethane, respectively, is illustrated in FIGURE 4.[6]

These respective mechanisms may be formulated, and expressions obtained for arriving at the constants, K_2 and K_3, from the data of experiments. Thus, letting I_2 represent the observed luminescence intensity in the presence of a given concentration of inhibitor, we have, for Type I,

$$I_2 = \frac{bk(LH_2)(A_0)}{1 + K_1 + K_2X^r + K_1K_2X^r}, \tag{5}$$

in which X represents the molar concentration of inhibitor, and r the ratio between the inhibitor and enzyme molecules in the complex formed. Dividing EQUATION 2 by EQUATION 5 and simplifying,

$$\left(\frac{I_1}{I_2} - 1\right) = K_2 X^r. \tag{6}$$

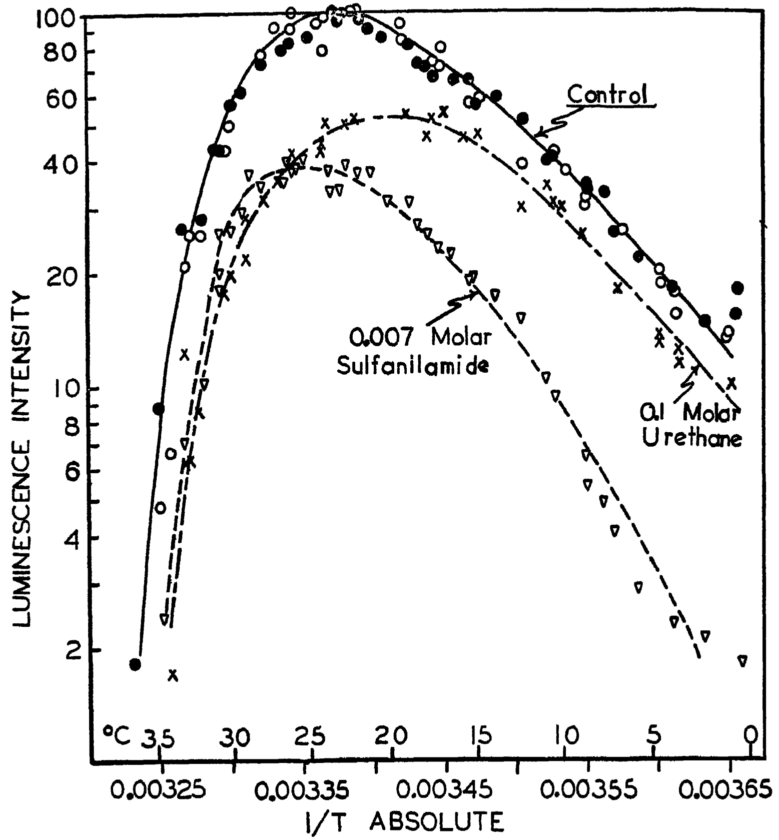


FIGURE 4. Luminescence intensity of *Photobacterium phosphoreum* in relation to temperature in a control and in corresponding suspensions containing 0.007 molar sulfanilamide and 0.1 molar urethane, respectively. The solid and hollow circles for the control are from repeated experiments. The data are replotted from Johnson, Eyring, Steblay, et al., 1945.[6]

Similarly, for Type II,

$$I_2 = \frac{bk(LH_2)(A_0)}{1 + K_1 + K_1K_3U^s}, \tag{7}$$

in which U represents the molar concentration of inhibitor, and s the ratio of drug-enzyme molecules in the complex. Dividing EQUATION 2 by EQUATION 7 and simplifying,

$$\left(\frac{I_1}{I_2} - 1\right) = \frac{K_1K_3U^s}{1 + K_1}, \text{or} \left[\left(\frac{I_1}{I_2} - 1\right)\left(1 + \frac{1}{K_1}\right)\right] = K_3U^s. \tag{8}$$

With the aid of EQUATIONS 6 and 8, the two mechanisms may be distinguished. In either case, a plot of $ln\left(\frac{I_1}{I_2} - 1\right)$ against ln molar concentration of inhibitor at constant temperature and pressure, yields straight lines whose slope gives the average number of inhibitor molecules combining with each enzyme molecule (FIGURE 5), unless, of course,

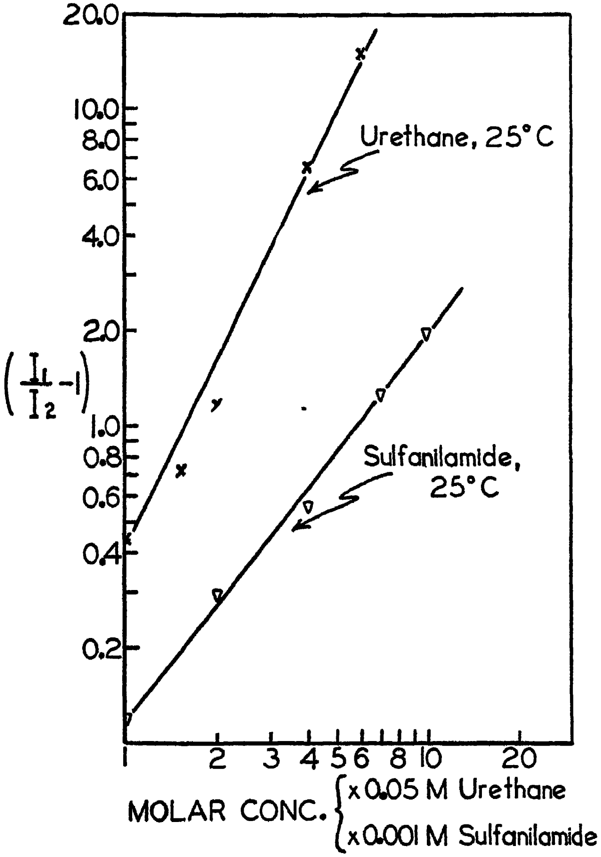


FIGURE 5. Relation between the concentration of sulfanilamide and urethane, respectively, and the amount of inhibition of bacterial luminescence, (*Photobacterium phosphoreum*) at 25° C., plotted in the manner discussed in the text.[6] The slope of the line for sulfanilamide is 1.2; for urethane, 2.0. With urethane, the slope increases with temperature, although with sulfanilamide there is little change.

more than one system is affected or additional reactions are involved, so that the action is more complicated than the theory takes into account. When, on the other hand, *ln* $\left(\frac{I_1}{I_2} = 1\right)$ is plotted against the reciprocal of the absolute temperature, a straight line over a wide range of temperatures is obtained for Type I, but not for Type II (FIGURE 6). From the slope of the straight line, the heat of reaction, ΔH_2, of the equilibrium may be computed. If the mechanism is Type II, a straight line,

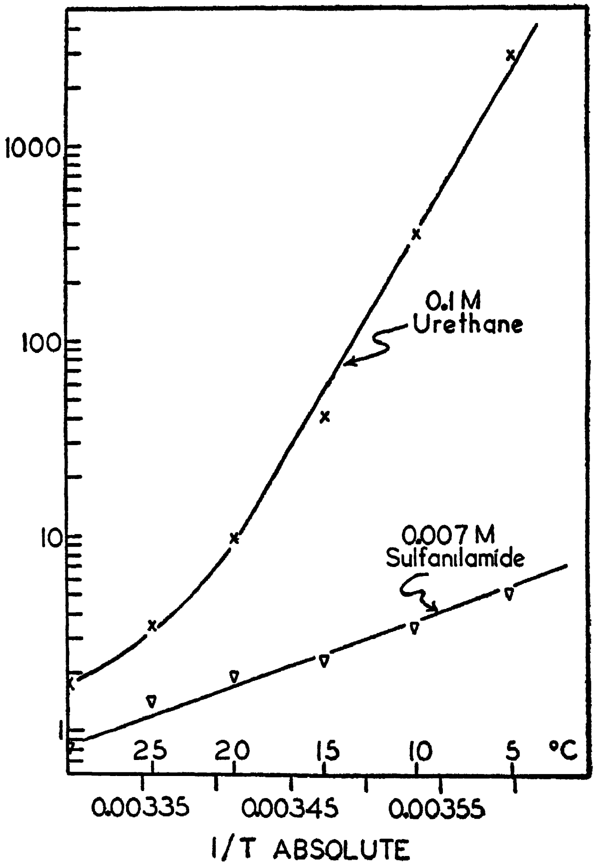


FIGURE 6. Analysis of the curves shown in FIGURE 4, according to the formulations given in the text. The luminescence intensity of the sulfanilamide-containing suspension (I_s) in relation to that of the control (I_1) is plotted on the logarithmic scale of the ordinate as $\left(\frac{I_1}{I_s}-1\right)$ against the reciprocal of the absolute temperature on the abscissa. The slope indicates 12,500 calories as the heat of reaction in the sulfanilamide-enzyme equilibrium. The values of the expression $\left(\frac{I_1}{I_s}-1\right)$ for the urethane curve in FIGURE 4 in relation to the control have been multiplied by $\left(1+\frac{1}{K_1}\right)$, and the product plotted on the ordinate for different values of $1/T$. In this case, the slope of the line, through the lower temperature range, indicates a heat of reaction of approximately 60,000 calories in the urethane-enzyme equilibrium.

whose slope depends on the heat of reaction, ΔH_s, results when $ln\left[\left(\frac{I_1}{I_s}-1\right)\left(1+\frac{1}{K_1}\right)\right]$ is plotted against $\frac{1}{T}$ (FIGURE 6). In the cases studied, however, the relation frequently departs from linearity, in the direction of too high an inhibition at temperatures beyond the normal optimum. Inasmuch as such deviations have been found to be more marked with increasing concentration of the drug, and since Type II inhibitors have been found to promote an irreversible as well as a reversible denaturation of the luminescent system,[6] it appears likely that, under these conditions, the irreversible effects are added to the reversible ones. Only the latter are taken into account by the above formulation.

In a purely diagrammatic way, the influence of temperature, pressure, and inhibitors of the two types discussed above may be illustrated as shown in FIGURE 7. Thus, at given concentrations of substrate and enzyme, the rate of the reaction in a single system, or in a complex process consistently limited largely by a single system, will be accelerated by a rise in temperature, and retarded by an increase in pressure, in a manner quantitatively determined by the specific reaction rate constant, k_1. The extent to which the rate increases for a given rise in temperature depends on the heat of activation, $\Delta H_1^{\ddagger}$, in the equilibrium, $K_1^{\ddagger}$, between the reactants and the activated complex, which, once formed, decomposes with the universal frequency, $\frac{kT}{h}$, characteristic of all chemical reactions. Likewise, the amount of the pressure effect depends on the volume change of activation, $\Delta V_1^{\ddagger}$, in the same equilibrium constant, $K_1^{\ddagger}$. As long as the enzyme molecules remain within the active channel, enclosed

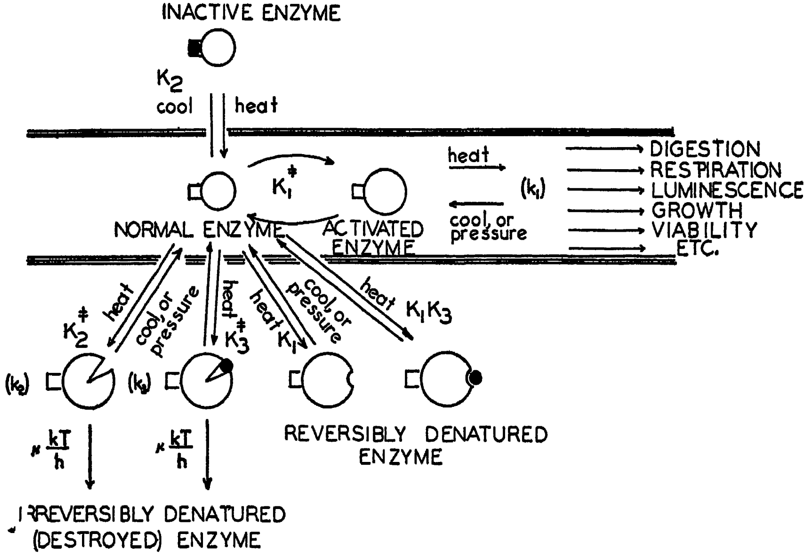


FIGURE 7. Schematic diagram to illustrate the control of the luminescent oxidation (k_1) in relation to temperature, hydrostatic pressure, reversible Type I inhibitors (K_2), reversible Type II inhibitors (K_3), the reversible temperature inactivation of the enzyme (K_1), irreversible thermal destruction (k_2), and irreversible destruction promoted by Type II inhibitors (k_3), as discussed in the text.

by the parallel lines (FIGURE 7), while the substrate concentration also is maintained at a constant, steady-state level and other factors remain the same, the reaction rate depends only on the value of k_1.

The amount of active enzyme, however, may be altered by one or more reactions. In the first place, as indicated above, it is apparent that the enzyme normally exists in equilibrium (K_1) between active and inactive forms. Because of the high heat and entropy of this reaction, the value of K_1 at temperatures below the optimum is so small as to be negligible. In the region of the optimum and above, K_1 increases rapidly with rise in temperature, causing the proportion of inactive molecules to increase to a greater extent, by a given increment in temperature, than the activation process in the enzyme reaction is accelerated: *i.e.*, K_1 is more strongly influenced by temperature than k_1. As a result, the observed rate goes through a maximum. Similarly with pressure: the inactivation proceeds with a large volume increase and, as a consequence, K_1 is decreased with rise in pressure and is more markedly affected than k_1, other factors remaining the same.

In addition to the reversible inactivation, with an equilibrium constant indicative of a protein denaturation equilibrium, a rate process of thermal destruction also takes place. In luminescence, the latter reaction has an even higher activation heat and entropy than the ΔH and ΔS of the reversible denaturation (FIGURE 8).[6] It is for this reason that the reversible denaturation of the luminescent enzyme can be so readily observed experimentally. The thermal destruction also proceeds with a large volume increase of activation, and consequently is markedly retarded by pressure at a given temperature (FIGURE 9).[6]

Both the reversible and irreversible denaturations are furthered by the addition of inhibitors of Type II, such as urethane or alcohol, apparently by the combination of these agents with bonds made available in the denaturation reaction. Both are opposed by hydrostatic pressure or by cooling.

By a mechanism different from those just described, certain inhibitors (Type I) enter into an equilibrium, K_2, causing an inactivation of the enzyme independently of the denaturation. The combination apparently occurs equally with the active and reversibly denatured forms and does not itself lead to an alteration in structure of the protein accompanied by a large volume change. Consequently, pressure has little influence. On the other hand, the enzyme-inhibitor complex will, in general, be dissociated by a rise in temperature, so that the fraction of inactive molecules will be less at higher than at lower temperatures, and the per cent inhibition correspondingly smaller.

The generality of the mechanisms, as shown in FIGURE 7, to a large extent awaits justification by further experiments, inasmuch as data sufficient for analysis on the basis of the theory described are not widespread,

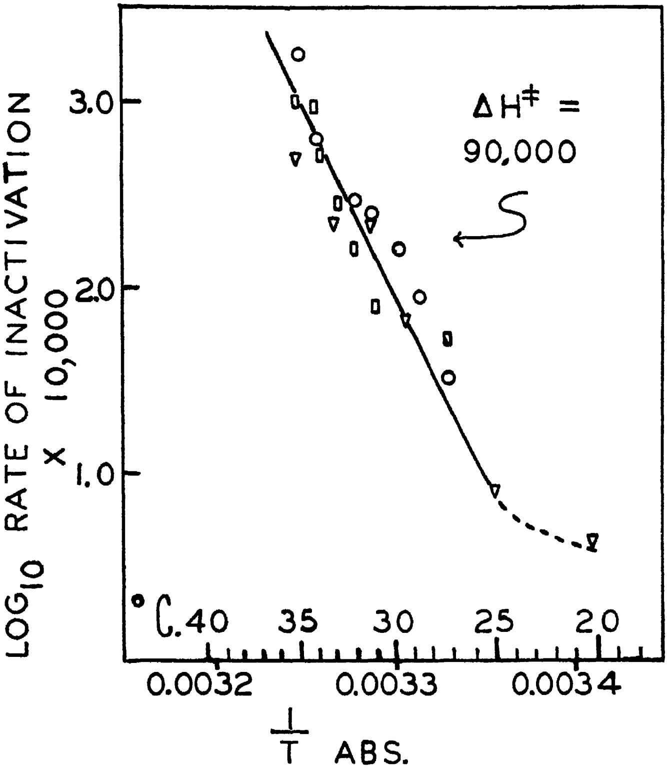


FIGURE 8. Temperature analysis of the rate of thermal destruction of the luminescent system in bacteria (*Photobacterium phosphoreum*). The points are from three repeated experiments. The slope of the line indicates an apparent activation energy of some 90,000 calories.[6] Semi-log scale.

particularly with respect to the action of hydrostatic pressure. Moreover, previous to the studies in connection with luminescence, the possibility that protein denaturation might be opposed by hydrostatic pressure had apparently not been taken into account. The opposite effect has been known for some years, *viz.,* that very high pressures, of the order of 10,000 atmospheres, at room temperature denature proteins, kill microorganisms, and inactivate enzymes, viruses, antibodies, bateriophage, etc. (*cf.* reviews by Macheboeuf and Basset[23] (1934), Cattell[24] (1936), and Bridgman[25] (1946). It is perhaps worth while, therefore, briefly to consider some of the available evidence with respect to other systems than luminescence.

First of all, it is apparent that the reactions by which proteins are denatured at room temperature under very high pressures are not identical with the reactions which take place at relatively high temperatures (or at lower temperatures in the presence of certain drugs such as alcohol) and which are opposed by moderate pressures, of the order of 500 atmospheres.

The fact that the lower pressures may greatly retard the denaturation of a highly purified protein, human serum globulin, at 65° C. has recently been demonstrated.[26] Small concentrations of ethyl alcohol accelerate the precipitation, while pressure retards it, with alcohol as well as without it (FIGURE 10). The rate is higher than first order. Preliminary estimates of the volume change indicate a volume increase of activation of about

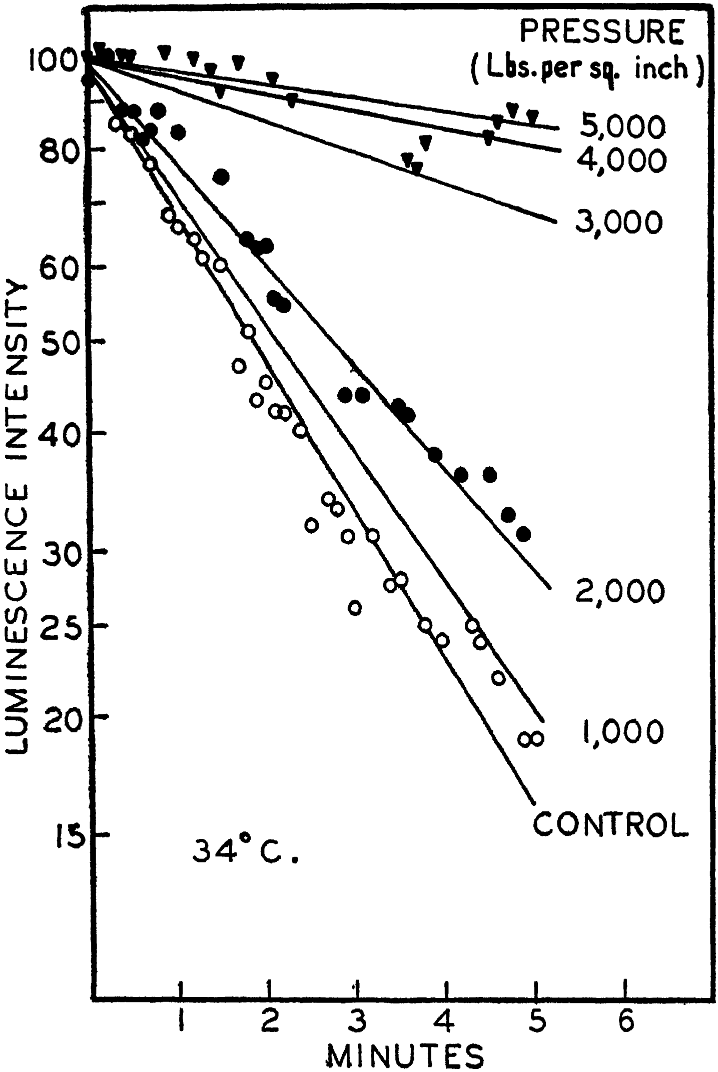


FIGURE 9. Influence of hydrostatic pressure in retarding the thermal destruction of the luminescent system at 32° C.[6]

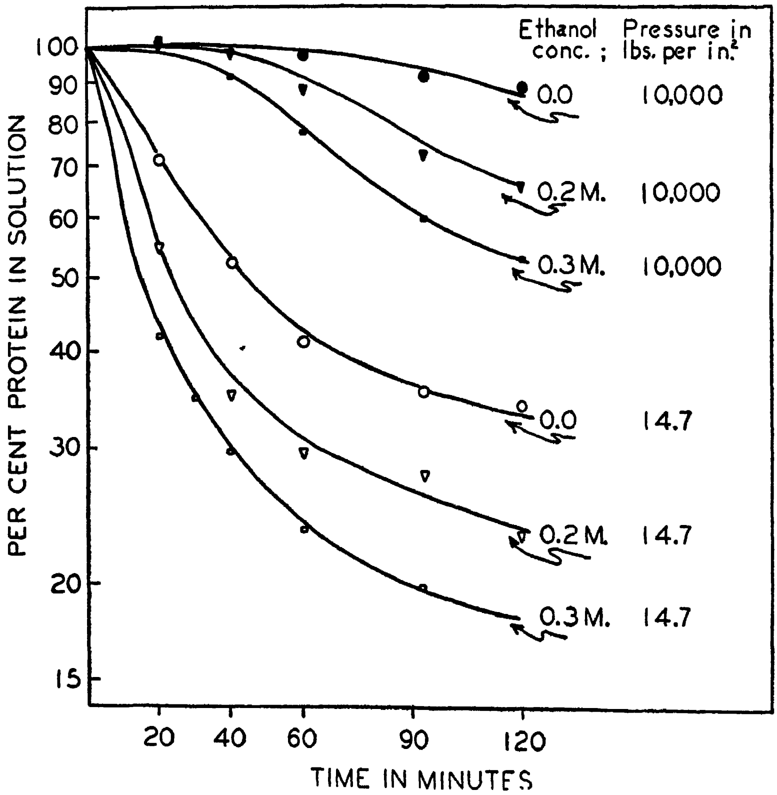


FIGURE 10. The rate of precipitation of highly purified human serum globulin at 65° C. in relation to hydrostatic pressure and small concentrations of ethyl alcohol.[26] Semi-log scale.

100 cc. per mole. The denaturation of specific antitoxic activity at the same temperature is also opposed by a pressure of 680 atmospheres.[27] Furthermore, specific precipitation of rabbit immune serum by a simple trihaptenic dye antigen, at room temperature, is greatly retarded under pressure of 680 atmospheres, indicating a large volume increase, of some 50 cc. per mole according to available data, in the reaction involving the antibody molecules.[28]

In regard to extracted enzyme and sol-gel systems, Marsland and Brown[29] have shown that the sol-gel equilibrium of rabbit myosin is characterized by a volume change of 120 cc. per mole. This is especially interesting, since sol-gel reactions influencing intracellular processes involving protoplasmic streaming, *e.g.* cyclosis, amoeboid motion, cleavage of animal cells, etc.,[30] are also accompanied by large molecular volume changes, of the order of 102 cc. per mole.[29] The rate of various extracted enzyme systems, *e.g.*, lipase, pepsin, and pancreatic proteinases (Benthaus, 1942[31]; *cf.* also Deuticke and Harren, 1938[32]), at room temperature, is reversibly retarded by pressures up to 1,500 atmospheres. As yet, sufficient data do not exist for a satisfactory analysis of the volume changes in these enzyme reactions. Thus, further studies should yield interesting results. Moreover, it might be expected that, under conditions favoring a reversible denaturation of the enzyme, an increase in the net activity of the system under pressure would be encountered, as in luminescence.

The activity of invertase or of diastatic enzymes is apparently not greatly retarded under moderate pressure at room temperature (Regnard, 1884[38]) and may even be increased.[21, 31] Recent studies[21] have shown that the increase in rate of sucrose inversion by invertase is most marked under conditions of partial inactivation of the enzyme, such as high temperatures at optimum pH, or at lower temperatures in relatively alkaline or acid solution (FIGURE 11). Analysis of the data indicated that at pH 7.05 and 35 or 40° C., the enzyme undergoes a reversible denaturation with a volume increase of about 69 cc. per mole.

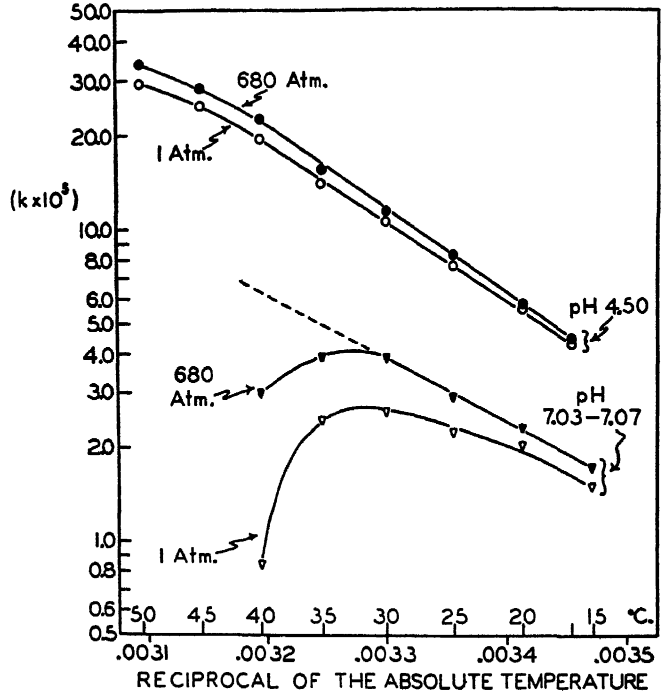


FIGURE 11. The rate, during the logarithmic period, of inversion of 10 per cent sucrose by invertase, in relation to temperature and hydrostatic pressure, at pH 4.5 and pH 7.03-7.07, respectively.[21]

In living cells, a number of phenomena, such as the tension of auricle muscle at room temperature, increase in intensity under pressures up to about 400 atmospheres, and then decrease as the pressure is raised up to less than 1,000 atmospheres.[34] The similarity in effects of pressure on these processes and on luminescence suggests that corresponding mechanisms are involved. In still other complex phenomena, including the rates of microbial growth and disinfection, a reversible inhibition by pressure has been noted.[35] A greater multiplicity of reactions are, no doubt, concerned in bringing about the measured result in all these cases, and the analysis, therefore, is more difficult than in luminescence.

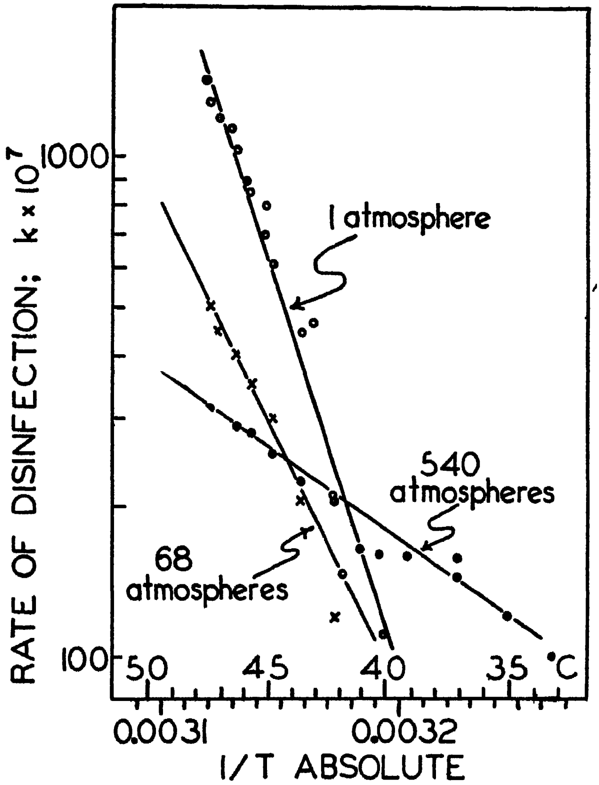


FIGURE 12. The rate of disinfection of non-proliferating cells of *Escherichia coli* in relation to temperature, under normal and increased hydrostatic pressures of 1,000 and 5,000 pounds per square inch, respectively.[35]

FIGURES 12 and 13 illustrate some of the data pertaining to growth and disinfection.

The relation between hydrostatic pressure and amount of inhibition caused by a given concentration of a drug has apparently been studied only in connection with bacterial luminescence,[6, 22, 36] and to a lesser extent, bacterial growth and disinfection.[35] Although the influence of

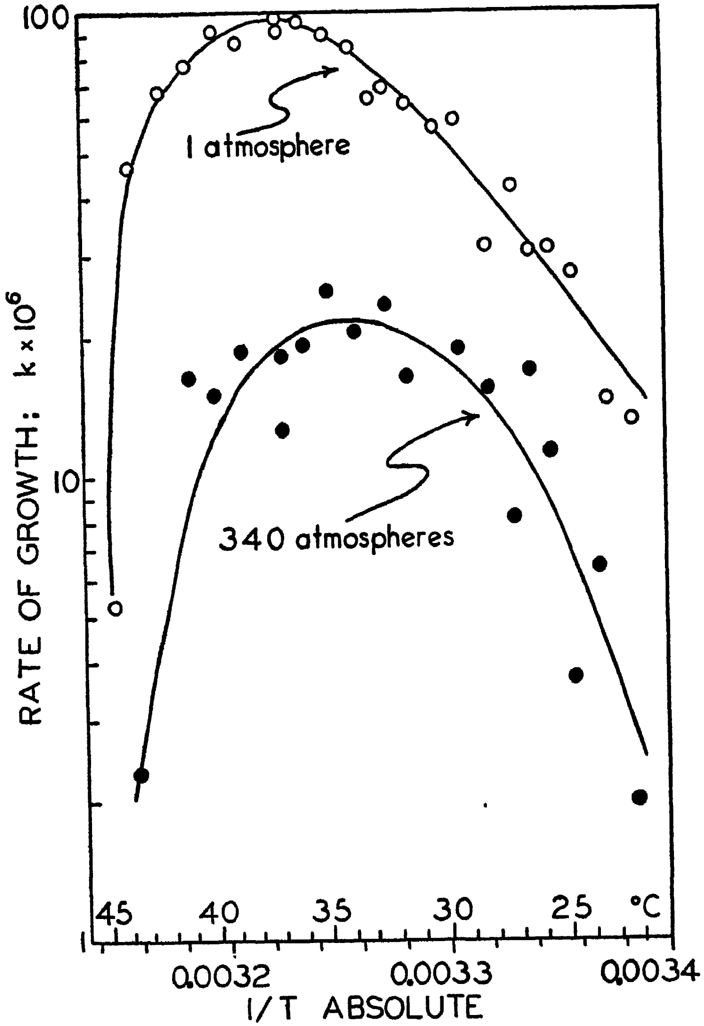


FIGURE 13. Growth (reproductive) rate of *Escherichia coli*, during the logarithmic phase, in relation to temperature, under normal and increased hydrostatic pressure of 5,000 pounds per square inch.[35]

temperature, at normal pressure, has received considerable attention, the data are not often sufficiently extensive, nor the mechanism of the reactions sufficiently well understood, to justify undertaking an analysis on the basis of the theory described above. According to the results of a recent investigation,[37] however, it appears that the rates of oxygen consumption and of methylene blue reduction by *Rhizobium trifolii* are affected by urethane in a manner resembling the action of this drug on bacterial luminescence. Using the same formulations, the quantitative relation between various concentrations of the drug, and the amount of inhibition at various temperatures, may be accounted for with some accuracy. The influence of pressure in this case has not been studied, but an analysis of its action should assist in elucidating the mechanism of the inhibition.

The evidence that large volume changes of reaction or of activation, respectively, take place in such diverse phenomena as luminescence, extracted enzyme reactions, cell division, protein denaturation, specific precipitation, the action of certain inhibitors, and so forth, would seem to justify more extensive studies from the point of view of hydrostatic pressure as well as temperature. Furthermore, there is considerable reason to believe that the synthesis of complex molecules, biologically specific in structure, involves a templet mechanism, and in order for a molecule to act as a templet it must be in a one- or at most two-dimensional form. This means that globular molecules would have to unfold before functioning as a templet. Such unfolding might be expected to be accompanied by fairly large volume changes. Studies with hydrostatic pressure should yield significant data in this connection also, and the theory worked out with luminescence as an indicator of protein reactivity will perhaps be found useful in various connections.

BIBLIOGRAPHY

1. **Eyring, H.**
   1935. J. Chem. Phys. **3**: 107.
2. **Glasstone, S., K. J. Laidler, & H. Eyring**
   1941. The Theory of Rate Processes. McGraw Hill. New York.
3. **Harvey, E. N.**
   1935. Erg. Enzymforsch. **4**: 365.
   1940. Living Light. Princeton University Press. Princeton.
   1941. Ann. Rev. Biochem. **10**: 531.
4. **Harvey, E. N.**
   1941. In: **Baumann, E., & K. Myrback.** Die Methoden der Fermentforschung. G. Thieme. Leipzig.
5. **Anderson, R. S.**
   1936. J. Cell. & Comp. Physiol. **8**: 261.
6. **Johnson, F. H., H. Eyring, R. Steblay, H. Chaplin, C. Huber, & G. Gherardi**
   1945. J. Gen. Physiol. **28**: 463.
7. **Chance, B., E. N. Harvey, F. H. Johnson, & G. Millikan**
   1940. J. Cell. & Comp. Physiol. **15**: 195.
8. **Eymers, J. G., & K. L. van Schouwenburg**
   1937. Enzymologia **3**: 235.
9. **Chase, A. M.**
   1943. J. Biol. Chem. **150**: 433.
10. **Harvey, E. N.**
    1928. J. Biol. Chem. **78**: 369.
11. **Van Schouwenburg, K. L.**
    1938. On Respiration and Light Emission in Luminous Bacteria. Thesis. Delft, Holland.

12. **Johnson, F. H.**
1939. Enzymologia **7**: 72.
13. **Johnson, F. H., K. L. van Schouwenburg, & A. van der Burg**
1939. Enzymologia **9**: 195.
14. **Anderson, R. S.**
1937. J. Am. Chem. Soc. **59**: 2115.
15. **Johnson, F. H., & A. M. Chase**
1942. J. Cell. & Comp. Physiol. **19**: 151.
16. **Johnson, F. H., D. E. Brown, & D. A. Marsland**
1942. Science **95**: 200.
17. **Brown, D. E., F. H. Johnson, & D. A. Marsland**
1942. J. Cell. & Comp. Physiol. **20**: 151.
18. **Johnson, F. H., H. Eyring, & R. W. Williams**
1942. J. Cell. & Comp. Physiol. **20**: 247.
19. **Eyring, H., & J. L. Magee**
1942. J. Cell. & Comp. Physiol. **20**: 169.
20. **Eyring, H., & A. E. Stearn**
1939. Chem. Rev. **24**: 253.
21. **Eyring, H., F. H. Johnson, & R. L. Gensler**
1946. J. Phys. Chem. **50**: 453.
22. **Johnson, F. H., D. E. Brown, & D. A. Marsland**
1942. J. Cell. & Comp. Physiol. **20**: 269.
23. **Macheboeuf, M. A., & J. Basset**
1934. Erg. Enzymforsch. **3**: 303.
24. **Cattell, McK.**
1936. Biol. Rev. Proc. Camb. Phil. Soc. **11**: 441.
25. **Bridgman, P. W.**
1946. Rev. Mod. Phys. **18**: 1.
26. **Johnson, F. H., & D. H. Campbell**
1945. J. Cell. & Comp. Physiol. **26**: 43.
1946. J. Biol. Chem. **163**: 689.
27. **Johnson, F. H., & G. G. Wright**
1946. Proc. Nat. Acad. Sci. **32**: 21.
28. **Campbell, D. H., & F. H. Johnson**
1946. J. Am. Chem. Soc. **68**: 725.
29. **Marsland, D. A., & D. E. S. Brown**
1942. J. Cell. & Comp. Physiol. **20**: 295.
30. **Marsland, D. A.**
1942. Protoplasmic Streaming in Relation to Gel Structure in the Cytoplasm. In: *The Structure of Protoplasm.* Collegiate Press. Ames, Iowa.
31. **Benthaus, J.**
1942. Biochem. Z. **311**: 108.
32. **Deuticke, H. J., & P. Harren**
1938. Z. Physiol. Chem. **256**: 169.
33. **Regnard, P.**
1844. C. R. Soc. Biol. **36**: 164.
34. **Edwards, D. J., & D. E. S. Brown**
1934. J. Cell. & Comp. Physiol. **5**: 1.
35. **Johnson, F. H., & I. Lewin**
1946. J. Cell. & Comp. Physiol. **28**: 23, 47, 77.
36. **Johnson, F. H., & L. Schneyer**
1944. Am. J. Trop. Med. **24**: 163.
37. **Koffler, H., F. H. Johnson, & P. W. Wilson**
1947. J. Am. Chem. Soc. **69**: 1113.

THE ANATOMY AND PHYSIOLOGY OF THE LIGHT ORGAN IN FIREFLIES*

By John B. Buck
National Institute of Health, Bethesda, Maryland

TABLE OF CONTENTS

I. Introduction 398
II. The Anatomy of the Light Organ 399
  A. Gross anatomy
  B. Histological types of light organ structure
  C. The "reflector" or "urate" layer
  D. The tracheal end-cell
  E. The photogenic cytoplasm
  F. The tracheoles
  G. The nerves
  H. A possible ultra-tracheolar network
III. Physiological Aspects of Luminescence 419
  A. Introduction
  B. Normal types of light-emission
  C. Localization of luminescence
  D. Tracheolar properties of interest in regard to control of luminescence
  E. The nervous system and luminescence
IV. The Control of Luminescence 432
  A. Historical
  B. Deductions from kinetics analysis
  C. Mechanisms of intracellular control
  D. Water as a possible limiting factor in the control of luminescence
  E. Luciferin and luciferase as possible limiting factors in the control of luminescence
  F. Oxygen as a possible limiting factor in the control of luminescence
  G. The spiracular factor in luminescence control
  H. The internal control of glowing
  I. Theoretical mechanisms of luminescence control by oxygen limitation
    1. Introduction
    2. Competition for oxygen, in relation to control
    3. The theory of osmotic control of luminescence
    4. The mechanical valve theory of end-cell control of oxygen
V. Summary 459
VI. Bibliography 463
VII. Figures 471

---

* The histological and cytological work which forms the basis for the original photomicrographs, and for various critical opinions appearing in this paper, was begun in 1938 at the Department of Embryology, Carnegie Institution of Washington, Baltimore, Md., and was continued at the University of Rochester during the period of 1939-1945. The author acknowledges with pleasure the technical assistance of Mrs. Gustav Kuerti, Miss Margaret Ramsay, Miss Olive Pitkin, and Miss Margaret Keister in the preparation of some two thousand microscope slide preparations, part of which was made possible by a grant-in-aid from the Society of the Sigma Xi. Grateful acknowledgment is likewise due to the National Research Council, and to the American Philosophical Society for grants which made possible the author's participation in the 1936 and 1941 expeditions to Jamaica under the auspices of The Johns Hopkins University. Thanks are due also to Miss Inez Demonet and associates for preparing the drawings for this report, and particularly to Dr. Rubert Anderson, Dr. Leigh Chadwick, Dr. E. N. Harvey, Dr. H. Specht, and Dr. Carroll Williams for extensive and helpful suggestions. Appreciation is due to the University of Rochester for generous help with the costs. Finally the author is enormously indebted to Elizabeth Mast Buck for aid with the manuscript.

INTRODUCTION

Since the days of Aristotle and Pliny, and presumably long before recorded scientific observation, the mystery of organic light has aroused the wonder of mankind. Fireflies being by far the most easily accessible luminous organisms, it is understandable that they were probably the earliest and most frequently studied forms. They are also outstanding in their own right, as a subject for scientific observation and experimentation, for perhaps no other animals have luminous organs of such size, brilliance, intricate structure, and physiological complexity. Therefore, it is not surprising that in the past two or three centuries an enormous literature has grown up on various aspects of firefly and glowworm luminescence. It is interesting to see that the problem has proved irresistible not only to specialists on bioluminescence but to a surprising number of men illustrious in other fields, as, for example, Swammerdam, Spallanzani, Davy, Faraday, Humboldt, Darwin, and Pasteur. In this paper, I shall not attempt to cover the literature exhaustively, but rather to review critically the major contributions bearing on the circumscribed problem outlined below.

As our knowledge of living light has increased, and the type of research has shifted from the naturalistic to the quantitative, it has turned out, for one reason or another, that certain aspects of the problem are better investigated on luminous organisms other than the firefly. Thus the crustacean *Cypridina* and luminous bacteria, in particular, have been used in most modern work on the chemistry, kinetics, and enzymology of bioluminescence. For the study of the isolated systems, the very complexity of the firefly light organ is, in a sense, a hindrance. On the other hand, it is well to recall that much of the pioneer work on bioluminescence, even on *in vitro* systems, was done on fireflies. Spallanzani (1796) used the glowworm in one of the earliest demonstrations that luminescence is dependent upon water and oxygen. In the tropical beetle *Pyrophorus*, Dubois (1886) first distinguished the enzyme and substrate of animal luminescence. It was, finally, on fireflies that Coblentz (1912) completed what still stands as the most comprehensive work on the spectroscopy and photometry of light-emission in living organisms. Moreover, fireflies still form the preeminent material for the study of at least two aspects of bioluminescence: the intimate anatomy of photogenic organs, and the physiology of the control of luminescence. This presentation will concentrate on these two aspects.

Since the main task of this paper is the integration of the structural details of the firefly light organ with the various manifestations of light production, it may be helpful briefly to outline the problem from the theoretical side, so that we may have a clear idea of what sort of information will be needed. The most fundamental level of the problem is the chemistry of the luminescent reaction. We need not be concerned with this, except to know the reactants. The next level of the problem concerns the general anatomy of the region in which the luminescent reaction takes place. Ideally, this would involve knowing where the different reactants are formed or obtained, how they are brought to the scene, and where they are held pending the reaction, or stored if in excess. The final level is the occurrence and control of luminescence. For this, we will need to know how the reactants are brought together at the desired moment, exactly where the reaction occurs, what sorts of luminescences are possible, which of the reactants is made limiting in stopping the reaction, and what physiological mechanisms operate in this control. For the present, we must be content with very incomplete knowledge at each level.

A long series of researches, initiated by Dubois (1885-1919)* and given quantitative expression by Harvey and his associates (Harvey, 1920 and 1940), show that the photogenic reaction involves a minimum of four reactants: water, oxygen, the substrate luciferin (a substance of low molecular weight, probably phenolic), and the enzyme luciferase. For most purposes, it makes little difference whether the oxygen is actually used in the energy-liberating reaction or reactions or in the recovery process.

THE ANATOMY OF THE LIGHT ORGAN

Gross Anatomy. Even externally, the luminous organs of the firefly are so diverse that the only generalizations possible are that they are close to the body surface behind a window of translucent cuticle, and are usually different in male, female, and larva. In size, the organs vary from minute pin-heads to masses occupying the entire ventral surfaces of several abdominal segments. In outline they vary from circular to entirely irregular. In position, they are found from the head to the tip of the abdomen, including the thorax, and on both dorsal and ventral surfaces. A number of representative luminous beetles are shown in FIGURE 1. Most of our common fireflies *(Photinus* and *Photuris)* are of the sort in which the organs occupy sternites 6 and 7 in the male (FIGURE 1, *a*), variable portions of 6, or occasionally 6 and 7, in the female (FIGURE 1, *b* and *e*), and two small spots on the ventral surface of sternite 8 in the larva (FIGURE 1, *c*). This larval position is the same as that in the adults of both sexes of the tropical genus *Diphotus* (FIGURE 1, *d:* Barber, 1941), which fact raises evolutionary questions. Another interesting type is that illustrated by the common European glowworms *Lampyris noctiluca* (FIGURE 1, *r*) and *Lamprorhiza (Lampyris) splendidula*† (FIGURE 1, *o*),

---

* In this paper, reference is made only to **Dubois'** 1886 monograph. Citations of his many other papers can be found in **Mangold** (1910) and **Harvey** (1920).

† The taxonomy of fireflies is in a very confused state. Among the points bearing on the subject matter of this paper are the following: (1) Many fireflies, to which the generic name *Lampyris* was applied originally, have since been put in other genera and hence are cited under different names in the older and in the more recent literature (examples: *Lamprorhiza splendidula; Luciola italica,* sometimes put in *Phausis; Luciola lusitanica; Phosphaenus hemipterus.* (2) There exist both *Photinus marginellus* and *Photinus marginellatus.* (3) The famous tropical firefly *Pyrophorus* is an elaterid (click-beetle) and is in a different family from all the other "fireflies" (the beetle family Lampyridae). (4) Because of inadequate descriptions in the literature, **Barber** was unable to identify, in my collections, most of the lampyrid species previously reported from Jamaica. It is probable, therefore, that some of the species used by **Lund** (1911) were actually studied by me under new names given by **Barber** (1941). I have followed custom in using the spellings "*pennsylvanica*" and "*Lamprorhiza*" for the taxonomically correct "*pensylvanica*" and "*Lamprohiza.*"

on which much work has been done. In these species, the most striking luminosity is seen in the wingless female. Another and very different type of organ distribution is seen in the famous tropical elaterid *Pyrophorus* (FIGURE 1, *m,n*). Both sexes of this beetle have small circular organs on the dorsal posterior corners of the pronotum, which emit a green light when the insect is at rest or walking, and a large rectangular organ in a cleft on the anterior surface of the abdomen, which emits an orange light when the insect is in flight. Some of the most spectacular displays of luminescence are seen in the larvae and larviform females of species in which the adult male is almost or entirely non-luminous. Such a creature is *Phengodes* (FIGURE 1, *s*), which has 11 or 12 segmentally arranged sets of photogenic organs giving a bright green light, and also its South American relative *Phrixothrix* (Harvey, 1944 and 1945), which has, in addition, a red light in its head.

Histological Types of Light Organ Structure. Firefly light organs show an astonishing diversity of structure and can be classified in a number of arbitrary ways. The system used here is an extension of Dahlgren's (1917), based on the arrangement of the tracheae.*

Typically, the photogenic cells are grouped together in one or more compact localized masses with specific tracheal and nervous supplies. However, one exception is found in *Phengodes,* (Type 1), where the light is produced by loose independent giant cells, apparently without tracheae, and similar to or identical with the oenocytes which are widely distributed among and within insects (FIGURE 15; see also Buck, 1940, 1942, 1946a).

The next more complex type of light organ structure is Type 2, found, for example, in *Phrixothrix* (Buck, 1946a), in the lateral "tuberculate" organs of the female of *Lamprorhiza splendidula* (Wielowiejski, 1882; Bongardt, 1903), and the larva of *Phausis Delarouzeei*.(Bugnion, 1929). Here (FIGURE 16) the organs are small subspherical compact masses of polyhedral photogenic cells with typically granular cytoplasm, constant in location, and with a specific tracheal supply. This latter consists simply of progressively tapering and more numerous branches, which form a sort of root-system ramifying through the photogenic tissue (FIGURE 8).

The third type of organ is like Type 2, except for the presence of a second layer of cells on the inner surface of the photogenic tissue. This layer is the so-called "reflector layer". Since it will be discussed later, it will suffice at the moment to state that it is composed of cells differing sharply from the photogenic cells in appearance, chemical composition, and staining reaction. This type of organ is illustrated for the larva of *Photuris pennsylvanica* (FIGURE 17) and for an adult *Diphotus montanus* (FIGURE 19) and is also characteristic of the larvae of *Luciola cruciata* and *Pyrocoelia rufa* (Okada, 1935a and b; Hasama, 1942c), *P. analis*

---

* The profuseness of the tracheal supply external to the photogenic organ has been mentioned many times and is well shown by Geipel (1915) and Hess (1921).

(Hasama, 1942c), *P. fumosa* (Hasama, 1944b), the larval organs of *Lampyris noctiluca* (see Wielowiejski and Bongardt), *Lamprorhiza splendidula* and *Phosphaenus hemipterus* (see Bongardt). The often-studied organs of the adult (female) of *Lampyris noctiluca* are also apparently of Type 3 (Owsjannikow, 1868; Wielowiejski; Bongardt). In addition, it seems clear, from the work of Heinemann (1886), Dubois (1886 etc.), Geipel (1915), Dahlgren, and others, that both the thoracic and abdominal organs of *Pyrophorus* show the compact but unorganized photogenic and reflector layers characteristic of Type 3, although the much larger size of the organs and much greater thickness of the layers tends to obscure the relationship (FIGURE 18). *Pyrophorus* also agrees with the abovementioned forms in having the simple arborescent tracheal supply seen in Type 2, in which the ultimate tracheal capillaries or tracheoles terminate between the photogenic cells and form the last link in an uninterrupted conduction system leading from the outside (FIGURE 8). The vertical spaces in the photogenic layer shown in Dubois' figure (FIGURE 12) have been doubted by Lund (1911) and Williams (1916) and are probably an exaggerated portrayal of the tendency, noted also by Dahlgren, of the photogenic cells to align in columns. The spaces could be caused artificially by these columns shrinking away from the tracheae running between them (which Dubois failed to see).

All the preceding types of organs differ from those now to be described in that they lack "tracheal end-cells." These cells, which were first described in fireflies by Schultze (1865), are structures which occur at the points where small tracheal twigs narrow rather suddenly and then divide into two or more delicate thread-like tracheoles, or tracheal capillaries, which *appear* to lack the spirally thickened walls which are characteristic of the tracheae proper. These end-cells seem to differ considerably in different material, or according to various workers, as is seen in FIGURES 2, 3, and 11 from Bongardt, Geipel, and Dahlgren, respectively. Nevertheless, there seems to be general agreement that they are uninucleate and have protoplasmic processes or extensions which surround and accompany the tracheoles for varying distances toward or among the photogenic cells, and give the end-cell a stellate appearance, which has led to their being compared with ganglion cells (Kölliker, 1858; Schultze; Eimer, 1872). It also seems clear that the end-cell is the principal site of reduction of inspired osmium tetroxide ("osmic acid") vapor, and, indeed, most of our knowledge of the end-cell was obtained by the use of this reagent (FIGURES 2, 3, and 11). An analysis of the significance of this reduction, and additional details on the internal structure of end-cells, are best deferred. It is worth pointing out, however, that end-cells are widespread in insects, and are even said, by Wielowiejski and Geipel, to occur in non-photogenic tissues of fireflies.

Among organs containing tracheal end-cells, there are a number of different anatomical arrangements which can be classified into convenient, though arbitrary, types. The simplest of these (Type 4) is reported by Dahlgren to occur in "some forms of *Photuris*" (though it does not appear to occur in either *P. pennsylvanica* or *jamaicensis*). As shown in FIGURE 4, the tracheae run ventrally through the reflector layer, as usual. However, when they reach the photogenic tissue, instead of penetrating the latter they divide into several branches which run laterally at the interface between the two layers, terminate in end-cells, and then send tracheoles down into the photogenic tissue, which is only one cell thick.

Another type (Type 5, see FIGURE 5) is found, according to Dahlgren, in the Japanese *Luciola parva, L. vitticollis,* and other oriental forms. The tracheal supply arboresces among the photogenic cells as in Types 2 and 3, except that, at certain points where the repeated branching and tapering has reduced the tracheae to quite small tubes, end-cells occur and give off the usual tracheoles. Thus defined, it agrees fairly well with the descriptions of the adult organs of *Pyrocoelia rufa* (Hasama, 1942a) and *Luciola africana* (Geipel), and seems also to include the older descriptions of the adult organs of *Lamprorhiza splendidula* (Schultze, Wielowiejski, Bongardt). The organ of *Luciola cruciata* (Okada, 1935b) is also placed in Type 5 provisionally, although in this organ (FIGURE 9) the larger tracheae appear not to be in direct contact with the photogenic cells, and the arrangement of the end-cells and tracheoles in horizontal section is more regular than expected (FIGURE 9b).

The Type 6 arrangement (FIGURES 7 and 10) is very common, being found in all American members of *Photinus* and *Photuris* which have been studied (*Photinus marginellus*—Townsend, 1904; Dahlgren. *P. pyralis*—Seaman, 1891; McDermott and Crane, 1911. *P. consanguineus* —Williams; *Photuris pennsylvanica*—Seaman; Lund; McDermott and Crane; Williams; Hess, 1922); in some twenty species of Jamaican *Photinus* (Lund; Buck, 1940 and 1942); in *Luciola parvula* (Hasama, 1944b); *Luciola italica* (Tozzetti, 1870; Emery, 1884); and doubtless many others. It is desirable to describe this type in some detail, in order to establish an adequate morphological basis for the future discussions on physiology. As a rule, the Type 6 organs are large, and situated on the ventral surfaces of sternites 6 and 7 in the male and 6 in the female (FIGURE 1, *a,b*). The reflector layer of the organ is in the inner or dorsal position, while the photogenic layer is ventral (FIGURE 20).

In Type 6, the tracheal trunks which supply the light organ run vertically through the reflector layer as usual, but when they reach the photogenic layer they do not pass directly between the photogenic cells but into specialized cylindrical rods of tissue which go straight through the photogenic layer to its ventral surface (FIGURES 7, 10, 23-26, 31, and 32).

These "cylinders"* contain a number of structural elements in addition to their axial tracheal trunk. First, there are short tracheal "twigs" which branch from the vertical trunk. Second, there is the tracheal epithelium, which, though extremely thin, has fairly large nuclei both along the trunk and the twigs. Third, there are tracheal end-cells, one at the end of each twig, with a nucleus about the size of those of the tracheal epithelium. Fourth are the tracheoles or tracheal capillaries which issue from the end-cells at just about the periphery of the cylinder and run out into the photogenic tissue. Though the existence of a differentiated membrane surrounding the cylinder is not readily apparent in many forms, some preparations suggest such a structure (FIGURES 29 and 38). The integrity of the cylinder tissue as a morphological unit is shown by the behavior of isolated cylinders in maceration preparations from which the photogenic cells have been removed (FIGURES 27, 37, and 40). In addition to these structures, there presumably is some sort of fluid or gelatinous matrix filling the apparent spaces in the cylinder around the end-cells, twigs, etc. Strangely enough, this material does not stain with any technique yet devised, so that in sections of the photogenic layer it is easy to come to think of the clear space around the tracheal trunk as "empty", and of the cylinders as hollow tunnels penetrating the photogenic tissue. If these spaces were indeed gas-filled, our ideas of how the photogenic tissue is supplied with oxygen would need revision.

In horizontal section (*i.e.*, one parallel to the external surface of the photogenic layer), or in surface view, the Type 6 organ shows a beautiful and characteristic "rosette" pattern, which has been remarked by numerous workers (Emery, Townsend, Geipel, Okada, etc.; FIGURES 10, 13, 22, 29, and 36). The vertical tracheal trunks are spaced in regular triangular symmetry, and around each the photogenic cells are arranged like the petals of a flower. In most species, the photogenic cells front upon two contiguous cylinders and have a roughly prismatic or rhombohedral shape.

The Type 6 organ has a number of interesting variations. For example, in many species, the cylinder often flares out like a trumpet or an hourglass at one or both surfaces of the photogenic layer, and there accommodates especially rich brushes of tracheal twigs and end-cells (FIGURES 28, 30, and 39). Such a cylinder in *Photuris jamaicensis* appears in FIGURE 25 in longitudinal section and is well shown in ventral surface view by Okada (1935b). Lund described the photogenic layer of the female of *Photuris pennsylvanica* as a single layer of cells, through which the tracheae run directly without cylinders, and which is penetrated by tracheoles (from end-cells) only from the dorsal and ventral surfaces (FIGURE 14). Such a structure is often seen. In other regions, however, even when only one cell thick, conventional cylinders occur with end-cells from which tracheoles run laterally. Possibly, Dahlgren's photurid organ (Type 4) was based on a similar structure in which the end-cells at the ventral surface were overlooked.

---

* Called "digitiform acini" by Tozzetti and by Emery.

The arrangement in *Photinus pallens* is, in one sense, the reverse of that just discussed, in that its cylinder is commonly narrow at dorsal and ventral surfaces but expanded in the interior of the photogenic layer (FIGURE 26). Other species present numerous variations in the relative thickness of photogenic and reflector layers, number of cell layers in each, diameter of cylinders, etc. A few of those occurring in various species of *Photinus* are illustrated in FIGURES 23, 24, and 31. One of the most unusual of these is the organ of *Photinus evanescens montego*, which has an extraordinarily broad and short cylinder, with profusely branched tracheae. A high-power view of this organ shows well the terminations of the numerous tracheal twigs and the limits of the cylinder tissue (FIGURE 31).

One rather puzzling feature of the Type 6 organ concerns the relationship between the processes of the end-cells and the cylinder walls. In *Luciola italica* (Emery), *Photinus consanguineus* (Townsend, Dahlgren), *Luciola parvula* (Hasama, 1944b), and in all of the American and Jamaican species of *Photinus* which I have examined, the end-cells are apparently as figured by Townsend and Dahlgren, and seem to be wholly contained within the cylinder and to give off tracheoles of uniform diameter which penetrate the photogenic tissue alone. In *Photinus marginellatus*, however, which has been studied carefully by Geipel (FIGURES 3 and 6), in *Photuris pennsylvanica*, according to Lund, and in both *P. pennsylvanica* and *P. jamaicensis*, according to my observations, the end-cells appear to have quite thick tapering processes which enclose at least the proximal parts of the tracheoles (FIGURE 32). These processes actually penetrate varying distances into the photogenic tissue, even though the end-cell body itself may lie within the confines of the cylinder or in little bays in the photogenic tissue (FIGURE 14). The arrangement, therefore, resembles in some respects that described in *Lamprorhiza* (Type 5) by the earlier workers (FIGURE 2), although there is no doubt that both *Photinus marginellatus* and *Photuris pennsylvanica* have the vertical cylinders characteristic of Type 6. The matter will be considered further under the discussion of end-cells.

The "Reflector" or "Urate" Layer. The reflector layer differs from the photogenic layer in a number of respects. First, the cytoplasm of its cells usually fails to stain with ordinary acid dyes such as eosin, so that in conventional histological preparations the layer is clear, in contrast to the heavily stained photogenic layer. Second, its cells are packed with small, highly refractile particles, probably birefringent and often described as "crystals" or "sphaerocrystals," but also reported by some to have a rounded outline. In sections of fresh tissue, these particles render the reflector layer opaque by transmitted light, in contrast to the translucent photogenic layer, and white by reflected light, in contrast to the dark photogenic layer. In balsam preparations, on the other hand, the "crystals" are mainly dissolved out by the dilute alcohols and leave the reflector layer nearly hyaline. Other differences, such as solubility in various reagents, have been described (Kölliker, Schultze, Wielowiejski, Bongardt, Townsend; Hasama, 1942a). End-cells have been reported from the reflector layer by Geipel, but according to Lund, Bongardt, and Emery, are very rare or absent.

The two-layered nature of the light organ was first recognized by Kölliker, who described chemical tests purporting to show that the material in the dorsal (reflector) layer is ammonium urate. While there is no reason to doubt this conclusion, and although the work has been repeated a number of times with only minor discrepancies (Owsjannikow, 1868; Tozzetti, 1870; Heinemann, 1872 and 1873; Bongardt; Lund) it has not been checked by modern analytical techniques. In assessing all the work on the chemical composition of the luminous organ or its layers, it must be kept in mind that the two layers have not been separated before testing, so that even microchemical color reactions *in situ* are by no means devoid of the possibility of contamination from the contiguous layer, or indeed from other viscera.

No agreement exists as to the function or functions of the reflector layer. Most workers have accepted, explicitly or implicitly, Kölliker's idea that it serves as a physical reflector to increase the amount of light emitted. No experimental work has been done on this point, however, and in view of the granular type of cytoplasm, I doubt whether the reflecting efficiency of the layer can be very high. In addition, much light would be absorbed in passing back and forth through the photogenic layer. The somewhat similar notion (McDermott and Crane) that the layer may act as a screen to protect the deeper-lying tissues from the light seems likewise to be questionable, because of the lack of evidence or expectation that the light produced in the photogenic cells is injurious.

The presence of urate in the reflector layer has led to a flood of mostly gratuitous speculations on possible metabolic connections between photogenic and reflector layers, even including a contention that the urate granules themselves are the true source of light (Weitlaner, 1909). Although a membrane between photogenic and reflector layers has been described (Tozzetti), the great majority of accounts show the two layers abutting directly (or even with a contact line so irregular that isolated photogenic cells project into indentations in the reflector layer and *vice versa*), so that a direct passage of particulate material between them is at least theoretically possible.

One idea which had considerable vogue provides that the cells of the photogenic layer, as they get "used up" and filled with the waste products of their activity, are transferred to the reflector layer (Schultze, Wielowiejski,* Dubois; Gerretsen, 1922; denied, however, by Emery and by Lund). Such a process is difficult to visualize for any but the photogenic cells already contiguous to the reflector layer, and would be equivalent simply to shifting the boundary between the layers rather than to actual cell movement. This would require, first, that the photogenic cells closest to the reflector layer, in order to be the first to transform, indulge in the most intense activity, and second, that with time the reflector layer should increase in thickness at the expense of the photogenic layer. Support for either of these points is lacking in most recent investigations, although either proof or disproof would be very hard to obtain, if for no other reason than the great difficulty in assessing and comparing total light emissions. As a matter of fact, "transition" cells† between the two layers have been described (Kölliker; Hasama, 1942a‡; see also FIGURE 23), and Weitlaner (1911) reported an increase in the amount of urate in old as compared with young individuals. On the other hand, Emery found no difference between young and old animals, and Townsend, Lund, and Geipel saw no seasonal change in thickness of either layer. Heinemann (1872) also found no change in four weeks in *Pyrophorus*. Okada (1935a) reported a decrease in thickness of the photogenic layer with age, without change in the thickness of the reflector layer. Hasama, in several species, found no difference in size or distribution of photogenic granules from fireflies long in darkness compared with fireflies after a long period of flashing. Moreover, the presence of transition cells has been denied specifically by several writers (*e.g.*, Lund, Geipel). In view of the considerable variability in layer thickness in different individuals, and particularly in different regions of the same organ (even if care is taken to use only sections exactly perpendicular to the surface), all such work is suspect until confirmed on numbers of specimens adequate to demonstrate statistically significant differences.

A somewhat similar idea is that the urate granules represent waste products of the oxidation of the photogenic granules in the light cells, which have been transferred by an undescribed process to the reflector layer and there stored (Kölliker, Tozzetti, Lund). Lund has championed this view strongly. His claim that the "waste product" deposited in the dorsal layer is "allied to or identical with some of the split products of nucleic acid" is interesting in view of the recent demonstration that firefly light organs contain a relatively high proportion of flavin-adenine dinucleotide (Ball and Ramsdell, 1944).* However, not only is there no decisive evidence of an actual transfer of material between layers, but, as Okada (1935a) has pointed out, there are difficulties in explaining how a waste material, which (according to Lund) crystallizes in the photogenic cells, comes to be redissolved, transferred (there is no blood circulation directly between the two layers), and recrystallized in the cells of the urate layer. Furthermore, Lund's "striking and conclusive" proof of a "direct relation and actual tracing of products of decomposition resulting from photogenesis from their place of origin in the photogenic cells into the dorsal layer cells" consists of no more than the following observation: When a cross-section of an entire photogenic organ is viewed by reflected light, the dorsal (reflector) layer was seen to be packed with a dense white mass (crystalline deposit), and in the ventral layer similar ("same") material was seen in the most peripheral cytoplasm and sometimes around the nuclei of the photogenic cells. Small amounts of this material were also observed by Lund in other viscera, and other investigators have reported that "urate" occurs in the fat body in considerable concentration. Lund also claimed that he found "different amounts of the accumulated products of katabolism in different species and different specimens of the same species," and that "the degree of filling of the dorsal layer cells also corresponds to the amounts of the deposit upon and in the photogenic cells." These claims are in direct opposition to the observations of several workers, particularly Emery, and Geipel. By implication, Lund regarded these differences as progressive stages in a process starting with no crystalline waste in the photogenic cells at the beginning of adult life, and culminating in a heavy deposit in aged specimens.

---

* Wielowiejski, in a later paper, came to the opposite conclusion in regard to American fireflies.

† Not to be confused with the term, transition cell, used by Holmgren (1895) and Townsend as synonymous for end-cell.

‡ Bongardt saw cells intermediate in character, but not at the boundary between the layers.

Lund's theory implies that at any moment the photogenic cells of a given firefly would present a cytoplasmic appearance directly related to the total amount of light which had been produced. Since such an assumption is prerequisite for all ideas of photogenic-urate transformation, it is germane to point out, first, that with a number of techniques a considerable cytoplasmic variation is apparent not only between the photogenic cells of the same individual but between contiguous cells, and second, that the finding of differences does not necessarily justify arranging them artificially into an irreversible temporal sequence. It seems more reasonable to assume that the photogenic cells go through individual, and probably asynchronous, cycles of activity which are repeated a number of times during the functional life of the photogenic layer. Such a view

---

[*] It is also interesting in view of McElroy & Ballentine's (1914) claim that phosphate is liberated during the *in vitro* luminescent reaction of *Cypridina*. McDermott (1915) attempted to test Lund's hypothesis by comparing the soluble phosphorus and nitrogen content of a solution of dried and powdered firefly organ which was allowed to luminesce in the presence of peroxide, with that of one prevented from luminescing. He found slightly less of both phosphorus and nitrogen in the solution which had luminesced, and concluded that Lund's theory was unsound. It is not clear, however, that a change in phosphorus and nitrogen should have been anticipated in McDermott's experiment.

would agree with what occurs in many gland cells. At present, we can only speculate on what this "activity" is. However, the schemes of enzymatic oxidation and resynthesis of luciferin which have been developed in recent years furnish a reasonable pattern (see summary in Johnson *et al.,* 1945). It is true that the cytological differences between cells of the same organ could be explained by assuming that all the cells are in a stable equilibrium of oxidation (photogeny) and synthesis, but that the point of equilibrium (intensity of activity) differs in different individual cells and is reflected in their appearance. Critical information bearing on these various possibilities would be of great importance in questions of chemistry and physiology of the photogenic process and the mechanism of control, but none exists aside from scattered observations of spotty or localized glowing of the photogenic organ. These, however, indicate the possibility of differences in the responses of different cells.

The fact that organs of Types 1 and 2 operate without a reflector layer could be construed as indicating that it cannot have any essential connection with photogeny, although Wielowiejski argued that, since these organs are characteristically very small, their wastes could diffuse away directly.

Recent biochemical schemes (*e.g.,* McElroy and Ballentine, 1944), provide for the resynthesis of at least part of the oxidized luciferin. If this occurs, and if the necessary energy is supplied by reactions having gaseous or easily diffusible end-products, there need be no accumulation of organic waste products in the photogenic cell and no necessity for providing elaborate mechanisms for their removal. There is also the possibility that reactions which are not reversible *in vitro* (*e.g.,* the luminescent oxidation of *Cypridina* luciferin) are reversible *in vivo,* making possible an even greater conservation of material by cyclic reutilization.

On the whole, therefore, there has been no clear demonstration of what the function or functions of the reflector layer are, although the large bulk of the tissue, its close association with the photogenic layer, and its regular presence in many fireflies suggest that it must be of utility to the organism.

The Tracheal End-Cell. The end-cell has been the subject of much detailed study, since its strategic position at the beginning of the tracheoles at once suggests that it may be concerned in controlling the oxygen which, presumably, passes through the tracheoles into the photogenic cells. The end-cell is by no means a typical cell. Ordinary techniques often fail to reveal much more than its nucleus. The cytoplasm is best demonstrated by subjecting the intact animal to osmic acid vapor. This apparently diffuses unaffected through most of the tracheal system, but escapes into the cytoplasm in the region of the end-cell, where it is reduced to the black suboxide or metallic form. Progressively heavier doses delineate progressively more of the cytoplasm and processes of the end-cell, and sometimes the tracheoles. The number of processes present usually corresponds to the number of tracheoles given off, where the two are distinguishable. It varies from two to seven in different species, though it is usually rather constant intraspecifically.

It is sometimes not clear how the tracheoles originate within the end-cell. Geipel shows the tracheal twig terminating blindly within the end-cell, without any tracheoles being present (FIGURE 3). However, the most popular idea involves a continuous tubular connection, within the end-cell, between the narrowed tip of the tracheal twig (or a common tracheole issuing from this tip), and the tracheoles (see FIGURES 2, 11, and 13; also the drawings of Wielowiejski). According to Bongardt, and Dahlgren, there is a darkly staining annular swelling ("rounded body" or "ampulla") around the common tracheole (FIGURE 11, "S"), for which sphincter properties have been postulated (Dahlgren; Creighton, 1926; Snell, 1932; Alexander, 1943). Little differentiation has been observed in the cytoplasm of the end-cell, except by Dahlgren, who described "contractile radial fibers" (FIGURE 11). This observation is considered in detail in the last section of the present paper.

Lund seems to have regarded the permeability to osmic vapor, and the ability to bring about its reduction, as specific and exclusive properties of the tracheal end-cell and the tracheolar wall or membrane. On the basis of the temperature lability of the osmic reduction effect, he even ascribed it to the presence of a specific "reductase", which he assumed also to perform an analogous function in transferring oxygen in luminescence.* Wielowiejski had a somewhat similar idea. However, it has long been known that, with heavy doses, the area of reduced osmium spreads beyond the end-cell into the cytoplasm of the contiguous photogenic cells. Moreover, it was shown by Wielowiejski that, in Type 3 organs (which lack end-cells, and in which the tracheoles do not all originate at one point), the tracheoles are nevertheless permeable to osmic acid vapor and the surrounding photogenic cytoplasm is able to accomplish its reduction.† This was also shown clearly in *Pyrophorus* by Geipel.

---

* Remy (1925) found (not in fireflies) that the tracheal wall recolorizes indigo-white and, hence, has an "oxidative power." It is not clear, however, why this effect is not due merely to the proximity of gaseous oxygen. Aside from luciferase, no enzymes have been separated from photogenic tissue. Among other enzymes which have been postulated is Gerretsen's "photogenase", which presides over the formation of luciferin. Some suggestion of catalase activity is contained in Burge's (1916) claim that fresh firefly tissue liberated more oxygen from hydrogen peroxide than did an equal weight of bee tissue, and that the "luminous part" of a firefly liberated more than the non-luminous. The work, however, was poorly controlled.

† Wielowieiski distinguished clearly and categorically between Type 3 organs with their tree-like tracheal branching (such as those of *Lampyris* or the minute "tuberculate" organ of the female of *Lamprorhiza splendidula*) and the Type 5 organs with end-cells of *Lamprorhiza splendidula*. He did point out, however, that sometimes in Type 3 organs the point of origin of tracheoles, especially if more than one tracheole is involved, may simulate an end-cell due to the reduction of osmium there. This is particularly true since the thin tracheal epithelium often fans out at such loci. Geipel apparently observed similar structures in the reflector layer of *Photinus marginellatus*, which he interpreted as a different type of end-cell. Bongardt, working with the same materials as Wielowiejski, apparently duplicated the latter's observations very closely, but his presentation and discussion are so incredibly incoherent and contradictory (culminating on his page 25) that I found it impossible to make out whether he even believed in the existence of end-cells at all.

One of the least clear features of the end-cell is its external surface. This subject, though rather esoteric, is considered here in some detail, because it may help to clarify a large mass of confused literature and because it bears on certain properties of the end-cell and tracheoles which are of physiological interest. In macerated preparations (FIGURES 27, 37, and 40), the end-cells maintain their integrities as separate masses of protoplasm. However, their apparent limits seem often to depend on the degree of osmic impregnation (*e.g.*, Schultze, Townsend, Lund) and often no sharp membrane can be made out in either fixed or fresh preparations. This fact is illustrated in Townsend's and Dahlgren's figures, and I have observed it in most of the species of *Photinus* which I have studied (FIGURES 31, 37, 39, and 40). Emery was even led to consider all the contents of the cylinder as a sort of syncytium, and both Emery and Townsend regarded the end-cells as artifacts. It is clear from their figures, however, that they were dealing with the same structures which we now call end-cells. On the other hand, many investigators figure the end-cells as with a definite boundary, although they are vague about its nature (*e.g.*, FIGURES 2 and 3). Part of the confusion results from the fact that the tracheal end-cell is believed to be a derivative of the thin tracheal epithelium which covers all the tracheae, including the lateral twigs in the cylinder (Lund; Williams; Hess, 1922). Wielowiejski and Emery, for example, conceive of this epithelium as stretched over the flat fan-like furcation point of the tracheoles, like the web of a duck's foot, thus forming the "membrane" of the end-cell. This membrane, in turn, is produced as an "end-cell process", and invests each tracheole individually throughout its course (or, according to Bongardt and Williams, part of its course). Over most of this distance, the epithelium is indistinguishable from the tracheolar wall proper, because of its extreme thinness. On the other hand, in regions where the two can be separated visually, they can be differentiated also by the use of strong alkali, which dissolves the former, leaving the chitinous (*sic*) tracheole.

Other descriptions of end-cells differ in several respects from that of Wielowiejski. According to Bongardt, the end-cell processes accompany the tracheoles for long distances in the photogenic tissue, branch and anastomose richly on the surfaces of the cells, but eventually terminate and leave the individual unbranched tracheoles to run alone farther into the tissue. Geipel, on the other hand, seems to find no typical tracheoles at all in *Photinus marginellatus*, but only long, hollow, tapering end-cell processes, like the roots of a molar tooth, which embrace the photogenic cells (see also FIGURE 10).* Likewise, most of the species studied by Hasama have tapering end-cell processes. Here, however, a tracheole runs in each process up to, but not beyond, its tip. My observations on *Photuris*, which has end-cells apparently very similar to those of *Photinus*

---

* The tracheoles shown within the end-cell and processes of Dahlgren's reproduction (his FIGURE 19) of Geipel's figure (here reproduced as FIGURE 3) are not in Geipel's original.

marginellatus, lead me to believe that Geipel overlooked all, and Hasama part, of the tracheoles. In *Photuris,* these extraordinarily delicate tubules can be seen running from the tracheal twig into each end-cell process, and then issuing from the distal end of each process and proceeding farther into the photogenic tissue. This is visible in preparations made by several standard methods, and thus is unrelated to the vagaries of osmic acid penetration.

In summary, it seems justifiable to conclude that all the types of end-cells discussed are built on the same fundamental plan of multiple tracheolar branching from a tracheal twig.

There seem to be two main types: the *Photinus* type (with some exceptions), in which the end-cell is strictly limited to the cylinder, has an indistinct "cell membrane", and gives off usually two tracheoles of uniform diameter (FIGURES 11, 22, 26, 31, and 38); and the *Photuris-Lamprorhiza-Photinus marginellatus* type, in which the end-cell projects into the photogenic tissue, has a definite "cell-membrane", and gives off usually four to seven apparently tapering processes which accompany the tracheoles for varying distances into the photogenic tissue (FIGURES 6, 10, 14, and 32). There are numerous additional minor discrepancies between the various accounts of end-cells. Some of these are doubtless *bona fide* differences in material. However, there is probably considerable justification for attributing much of the contradiction to inferior optics, since the homogeneous immersion lens was not in general use until about 1885. This is the more reasonable in that the details in question are near the limits of microscopic resolution.

The Photogenic Cytoplasm. The strikingly granular contents of the photogenic cells have attracted the attention of numerous workers. The granules are minute, usually spherical and densely packed, and take acid or plasma dyes. Dahlgren claims that tissues fixed in boiling fixatives show that the granules are spherical in the male, rod-shaped in the female (FIGURE 11). Hasama (1942a) could not confirm the distinction in *Pyrocoelia rufa,* but says that Okada found about half the females in *P. consanguineus* differing from the males in the manner claimed by Dahlgren. In *Photinus pallens,* I found that granule shape is more or less constant in a given individual and may differ in different individuals, but is not characteristic of either sex.

Dubois regarded the photogenic granules ("vacuolides") as self-perpetuating entities comparable to mitochondria, and a similar concept is implied in Dahlgren and Kepner's term "photochondria". Vonwiller (1921) claimed that the granules stained like mitochondria, but Takagi (1934) showed that mitochondria could be demonstrated independently of the granules in *Luciola cruciata.* Moreover, Dubois was unable to culture granules isolated from several organisms.

Kuhnt (1907) suggested, by analogy with leguminous root-nodules, that the firefly light organ contains symbiotic bacteria. Pierantoni (1914), influenced by his studies on the symbiotic "mycetome" organs of Hemiptera, claimed to have cultured two kinds of bacteria from both the photogenic organ and the egg of *Lampyris*. These "bacteria" were described as distinct from the photogenic granules. Their figured appearance strongly suggests that they are mitochondria. Pierantoni did not make single-cell isolation cultures, and his evidence for the transmission of the "bacteria" by way of the egg is unconvincing. Furthermore, his cultures were not luminous. This is not necessarily a decisive objection in view of the known dependence of bacterial metabolism on culture conditions. Also, similar results have been obtained in various organisms, particularly fish, where the evidence for bacterial symbioticism is much more sound than it is in fireflies (see Harvey, 1940, pages 30 to 36). Vogel (1922) and Hasama (1942a) were unsuccessful in culturing anything from the light organs of fireflies, although Hasama's failure may have been influenced by his choice of 37° C. as an incubation temperature. Buchner originally (1914) supported Pierantoni's thesis, though he never, himself, worked on Lampyrids. In the second edition of Buchner's book (1930), however, the matter is left undecided.

Harvey and Hall (1929) demonstrated that the development and functioning of the adult light organ in *Photuris pennsylvanica* is unaffected by ablation of the larval organs. They concluded, therefore, that bacteria are not concerned with luminescence. They acknowledged, however, the possibility that bacteria in a non-luminous phase might have existed in other parts of the body and have contributed to the formation of the adult organ.

On the whole, although the photogenic granules and other cytoplasmic inclusions often resemble bacteria in form, size and staining, and although some organisms apparently do have organs (even luminous organs) which contain symbiotic bacteria, the evidence in regard to fireflies is so weak that a definite conclusion would be wholly gratuitous.

A number of workers (*e.g.*, Williams) have described the staining reactions of the photogenic granules with a variety of dyes. Since, however, most such reactions are highly unspecific in a chemical sense, the most that can be concluded is that the granules probably contain protein. The granules have been identified as "the" photogenic material (Dubois; Dahlgren and Kepner, 1908; Lund, McDermott, Dahlgren, Williams, etc.), or even specifically as "luciferase", "photogenin" or "luciferin", with very tenuous justification except that granules of some sort seem to be associated with light-production in all known animals. It is outside the scope of this paper to go into the chemistry of the photogenic layer in detail, but it is obviously relevant to mention that analyses of *whole* organs by Dubois, Harvey, Gerretsen (1922), McDermott, and others, have shown that luciferase and luciferin are present (or, at least, that two extracts can be prepared which luminesce when mixed). These are universally assumed to be localized in the photogenic layer. Considering the certainty that the photogenic layer contains a number of compounds of varying complexity, and in view of the probability of contamination already mentioned, it is not surprising that earlier analyses of the material of the "photogenic layer" yielded identifications as diverse as "albuminous" (Kölliker), "lecithin-like" (Lund), and "phosphatide with an aliphatic radical" (McDermott, 1911b).

Another interesting feature in the cytoplasm of Type 6 photogenic layers is the striking "differentiated zone" which surrounds the cylinders, (figures by Lund; Hess, 1922; Dahlgren; Williams; and Okada, 1935b). This zone is formed by a part of the peripheral cytoplasm of the photogenic cells in which the so-called photogenic granules are lacking, and only an extremely fine-grained and compact-looking cytoplasm can be seen (FIGURES 10, 11, 22, 26, and 31). The zone is of varying thickness in different species and in some has apparently not been recognized. That it differs chemically or physically from the rest of the cytoplasm, is indicated by its different response to a number of dyes. It is traversed, of course, by the tracheoles or by the end-cell processes as they pass from the cylinder into the luminescent tissue. Thus it is, by inference, the region where photogenic material might first come in contact with oxygen. According to most workers, this differentiated region is not found between different photogenic cells but only between the cells and the cylinder, or along the dorsal and ventral surfaces of the photogenic layer (FIGURE 26). In some of my preparations, however, it apparently also forms a very thin layer along the internal faces of the photogenic cells. The zone thus envelops the photogenic cells completely and "insulates" their interior cytoplasm. Possibly with this thought in mind, Dahlgren made the stimulating but unsupported proposal that this zone is impermeable to oxygen and serves as a protection against the "entrance of any oxygen that might come into the cells except through the tracheal capillaries." Aside from the difficulty of visualizing the mechanism of this impermeability, many firefly photogenic organs are apparently not equipped with such a protective layer and yet do not show uncontrolled luminescence. There may be some merit in the general concept, however, since Wigglesworth (1930), and others have shown that the walls of even large tracheae are permeable to oxygen. This suggests, in view of the low partial pressures of oxygen necessary to support luminescence (see page 442), that the quantity dissolved in the general body fluids might be sufficient to support luminescence, if there were no insulation. Also, as suggested on page 424, the differentiated layer might possibly have some bearing on observed differences in types of luminous emission (see also page 449).

The final peculiarity of the photogenic cytoplasm here considered is the unobtrusiveness of its limiting membrane. To be sure, the presence of sharp cell boundaries is described, or implied, by numerous workers (*e.g.*, FIGURES 6, 9, and 10; see also Hasama, 1944b), but in other instances tribute is paid to the fact that often no membrane of any sort has been observed (FIGURES 7, 11, 14, and 23). Lund, in fact, frankly regards the photogenic cytoplasm as a syncytium. This may be so in some species, but I hope to show later that in *Photuris pennsylvanica,* to which Lund referred particularly, cell membranes can be demonstrated. One reason why the membranes are difficult to find in Type 6 organs, is that they are concealed or obscured by the tracheoles. For example, in a horizontal section of the photogenic layer, the "lines" delineating the rosette pattern ordinarily appear single, and it is virtually impossible to decide, even with progressive focusing, whether one is looking at a tracheole or at a cell membrane in edge view (FIGURE 22). Even a double line would not be conclusive, since "the" membrane separating two photogenic cells is presumably double. Likewise, in a cross-section of the organ, some membranes are in face or quartered view, and hence invisible. Usually, it is only when a section is chosen so that an edge view of a membrane and a cross-sectional view of tracheoles are seen simultaneously, that one can distinguish the two with certainty. Figures of such views have been published by Bongardt, Townsend (and Geipel, for the end-cell processes), and examples from *Photuris pennsylvanica* and *Photinus pallens* are presented in FIGURES 35 and 36. In horizontal sections of the organ, cross-sections of tracheoles can usually only be seen close to the ventral (outer) surface, since it is only here that many of the tracheoles run vertically (FIGURE 36).

The Tracheoles. The origin of the tracheoles and their general arrangement should be clear from the discussions and figures already presented. Measurements of the dimensions of tracheoles (or, indeed, of any part of the luminous organ) are rare. The only figures for tracheolar diameter are Wielowiejski's and Lund's 1.3 and 1.1 microns, respectively, and for length, Townsend's value of 20 to 60 microns (estimated from her figures on the distances between cylinders in *Photinus marginellus*). The quoted dimensions are of the same order of magnitude as the averages obtained by the writer, but there are wide differences between different species, and the diameter measurements are probably subject to large errors as well as to variations caused by different techniques of preparation. In *Photuris,* the tracheoles issuing from the ends of the end-cell processes appear to be much less than a micron in diameter. In organs of Types 2, 3, and 5, where anastomosis probably does not occur, the tracheoles are much longer and by no means uniform in diameter at their origins. Moreover, they taper, so that diameters approximating or exceeding the resolving power of the microscope may be reached near their distal ends.

One hotly debated morphological point, which has physiological implications also, is whether the tracheoles actually penetrate the cytoplasm of the photogenic cells or are exclusively extracellular in their courses. Heinemann (1872) reported that the tracheoles impale the photogenic cells in *Pyrophorus,* though Robin and Laboulbène (1873), and Geipel described them as applied to the cell faces. Lund strongly espoused intracellular penetration in the several lampyrids he studied. However, the majority of workers on lampyrids have reported that the tracheoles (or end-cell processes) run only between or on the outside surfaces of the photogenic cells (Wielowiejski; Emery; Watasé, 1895; Bongardt; Townsend; Geipel; Dahlgren; Williams; Okada, 1935b; and Hasama). Both intracellular and extracellular tracheoles have been reported in other kinds of insects (Wigglesworth, 1930 and 1939). Lund's stand was primarily based on finding tracheoles close to nuclei in the same focal plane, mainly in *Photuris pennsylvanica.* I think this observation was a misinterpretation, due to the fact that in this species, the limits of the cells are poorly delineated and the cells are often irregular in shape, overlapping, and scarcely wider than the nucleus (FIGURE 35). At any rate, I have found no case of intracellular penetration of tracheoles in any of some twenty Jamaican and American species of *Photinus* and *Photuris.* Perhaps the most convincing evidence comes from the type of preparation already referred to, in which cell membranes and cross-sections of tracheoles both appear (FIGURES 35 and 36). Another type of evidence is seen in surface views of freshly extirpated light organs which have been dried sufficiently for air to enter the tracheoles. Here, the tracheoles follow a course similar to that figured by Townsend (FIGURE 13), which corresponds exactly to the intercellular interfaces seen typically in the rosette pattern of horizontal sections (FIGURES 29 and 38).

Another disputed subject is whether the tracheoles end free or anastomose with other tracheoles. This has physiological implications, since anastomosis might permit gas flow through the tracheoles, whereas transport to free endings would almost certainly be by diffusion. Anastomoses have not been reported in organs of Types 2 and 3. In *Lampyris* and *Lamprorhiza* the existence of "loops" was claimed by Kölliker, and denied by Schultze, but since tracheoles were unknown at that time, it is not clear what was meant. However, Wielowiejski found anastomoses only occasionally, and Bongardt states that they do not occur.* Five of the apparently Type 5 species discussed in Hasama's papers are likewise of the *Lampyris* type, since the end-cell processes (which are said to wholly contain the tracheoles) do not anastomose (*Pyrocoelia rufa* and *P. analis; Luciola lateralis, L. cruciata* and *L. gorhami*). In most Type 6 organs, profuse tracheolar anastomoses between contiguous cylinders are described (McDermott and Crane, Lund, Townsend, Williams, and Buck, 1942). Emery saw no anastomoses in *Luciola italica,* though his FIGURE 7 closely resembles most Type 6 organs (*e.g.*, FIGURES 13, 29, and 39). Tozzetti apparently saw some anastomosis *within* the cylinder in *L. italica,* although his descriptions and figures are not entirely clear. Okada (1935b), however, shows the end-cell processes ending free (FIGURE 10). On the whole, there appears to be a systematic difference between Types 5 and 6 in regard to anastomosis of tracheoles. This should be kept in mind during the discussion of flashing behavior in the two types.

---

* However, Bongardt described the tracheoles as running in the end-cell processes, and the latter as branching and anastomosing richly!

[*Note Added in Proof.* Dr. A. Glenn Richards recently made a number of electron micrographs of tracheoles from fresh light organs from males of *Photinus pyralis* which I supplied. The prints (3800 X) show isolated tracheoles of very uniform diameter (about 0.25 μ) and up to 25 μ long. In accordance with Richards's findings in many other insects, the tracheoles show clear spiral thickenings throughout, and appear to end blindly, without anastomosis. I have some doubts as to whether the tracheoles are actually those originating from end-cells in cylinders, because in several instances they appear to arise in groups of three, whereas in histological preparations the number is invariably two. Moreover, the advantages devolving from the increased resolution of electron optics are somewhat offset by the necessarily drastic method of preparation (maceration in water or weak alkali, teasing, drying *in vacuo,* and exposure to electron beams). Nevertheless, Richards's pictures caution against any dogmatic insistence on the existence of tracheolar anastomosis.]

Several writers have quoted Schultze as having shown the tubular nature of the tracheoles by gold chloride impregnation, forgetting that he never saw the true tracheoles. Wielowiejski, however, demonstrated the lumen by infiltration with dyed soap. It is now generally agreed both that they are hollow, and that the tracheolar wall is exceedingly thin. Both facts are illustrated in those figures of Bongardt and of Townsend which show cross-sections of tracheoles, and in FIGURE 36.

Though all agree that the tracheoles are hollow, there is less unanimity on what they contain, at least in life. Schultze, Tozzetti, Emery, and Lund claimed that in freshly dissected photogenic tissue the "tracheoles" (in some cases the finer tracheae must be meant) are nearly or entirely invisible, because they contain liquid. Tozzetti suggested that the tracheoles might carry part of the blood circulation. Wielowiejski and Townsend found that glycerin enters the air-filled tracheoles of a dried organ from the inside out (*i.e.*, distal-proximal direction), from which they concluded—irrelevantly, it seems to me—that the tracheoles are air-filled in life. No observations have been made on live fireflies, nor are any to be expected, in view of the thickness and delicacy of the photogenic tissue. On the other hand, in insects such as fleas, and mealworm and mosquite larvae, which, admittedly, have different anatomies, the situation in life is apparently quite variable and changes under various conditions (Wigglesworth, 1939). It is clear that, in some instances, the tracheoles are normally air-filled down to diameters of 0.5 micron and, since this is the order of magnitude of firefly tracheoles, there seems to be no reason why they might not be air-filled. As a matter of fact, unless the tracheoles were gas-filled throughout, intracellular penetration of tracheoles would confer no advantage, since the distribution of oxygen in the photogenic cytoplasm (if such is the purpose of the tracheoles) must ultimately occur by aqueous diffusion, which could only be hampered by a tracheolar wall. This matter will be further discussed in the section on physiology.

The Nerves. The least-known major anatomical feature of the photogenic organ is its nerve supply. In spite of contrary statements in the literature, I have not found the ganglia supplying the photogenic segments to be disproportionately large in comparison with those of similar non-luminous insects. Kölliker and Schultze were the first of several to report having seen nerves in the photogenic organ. With rare exceptions, however, these structures were demonstrated by reagents and techniques which are not recognized today as having diagnostic value in the identification of nervous tissue. Therefore, since even the standard neurological techniques are too often unpredictable, the possibility of artifact deserves serious consideration. This is the more pertinent in view of the great difficulty reported in demonstrating the "nerves" and especially in distinguishing their fine branches from tracheoles. Also, some very atypical structures have been described, as, *e.g.*, the "knobbed" and "multinucleate" fibers of Wielowiejski and Bongardt.* Wielowiejski devoted several pages to a description of connective tissue fibers (in non-photogenic tissue) which previous workers had misidentified as nerves, indicating that the techniques then in use were by no means reliable.

However, taking the reports at face value, the nerves generally follow the tracheal system rather closely (though Wielowiejski denies this) and are distributed in roughly the same fashion. There is much disagreement as to the ultimate terminations of the nerves, which are, of course, the regions of greatest interest. Tozzetti and Emery found no connection between nerves and any sort of cells in *Luciola.* Kölliker and Schultze could not trace the finer nerves to their ends in *Lamprorhiza,* though Schultze thought it likely that they innervated the photogenic cells. Wielowiejski found direct connections between nerves and the surfaces of photogenic cells in *Lamprorhiza,* as did Owsjannikow (1868) in *Lampyris.* Owsjannikow even described the nerve as penetrating to the nucleus, though Wielowiejski decried this idea. Geipel and Hasama (1942a) reported that in *Photinus marginellatus* and *Pyrocoelia rufa,* respectively, the finest nerves connect directly with the end-cells (to the nucleus, according to Geipel), and in *Lamprorhiza* and *Lampyris*, Bongardt found them connected with the exterior of both end-cells and photogenic cells.

---

* Lindemann (1868) described "nerves" running to the "light-balls" in *Lampyris*, but his descriptions and conclusions are so bizarre that I can only conclude that he had mistaken some other tissue for the photogenic organ.

There is, therefore, no general agreement on the details of the nervous supply of the photogenic organ. In my opinion, a really convincing answer to this problem will require a full investigation which is devoted to this one point and makes use of modern neurological techniques.

A Possible Ultra-tracheolar Network. In preparations made by silver nitrate impregnation, the tracheoles are heavily outlined in black by precipitated silver. In cross-section, they appear as little black circles strung on dotted black lines which are formed by a light deposit of silver on the photogenic cell membranes (FIGURE 36). Wherever the plane between two tracheoles (which means the plane of a cell membrane) is horizontal or nearly so (that is, parallel with the stage of the microscope), the tracheoles are seen to be knit together by a close-meshed network (FIGURES 33, and 34). This remarkable structure is only visible here and there at any one focus in a given field, in agreement with the rarity of instances where a flat region of cell membrane happens to lie parallel with the surface of the section, but it is often possible to trace the network over a relatively large area by careful focusing. The individual strands of the network are far too slender (of the order of 0.1-0.2 microns) to make it possible to say whether or not they are tubular. I have found them in *Photuris pennsylvanica* and *P. jamaicensis*, and in *Photinus pyralis* and *P. pallens*.

This network could be interpreted as a mesh of extraordinarily fine tubules binding the tracheoles together, in a manner analogous to the capillary networks between arterioles and venules. This might make a more satisfying picture, physiologically, than tracheoles alone, because such a network might permit quicker and more uniform distribution of oxygen to the photogenic cell, and thus reduce the partial pressure of oxygen necessary to support luminescence. However, in spite of the sharpness, relative orderliness, and wide distribution of this structure, I am not insisting on the above interpretation for the present, because metallic impregnation is such a notoriously capricious technique that we must reckon with the possibility of an artifact. Furthermore, in maceration preparations, and in those prepared with caustic, where the tracheoles are relatively free from other tissue, they do not behave as if held together by a network, nor do they show a rough outline such as might be left if the network had been torn off.

On the other hand, "ultra-tracheolar" offshoots of tracheoles have been figured from other insects, though not in a network (Wigglesworth, 1939, FIGURE 164).* The reason that the network appears only with the silver nitrate technique is undoubtedly that the "tubules" are so delicate that they would be invisible without a completely opaque coating. This probably explains, also, why they have not been reported previously, since silver nitrate has been used on firefly organs only very rarely (Owsjannikow, 1868, Tozzetti; Geipel).

---

*The networks described by Wistinghausen (1895) and Holmgren are larger than those discussed here by a factor of ten. Other supposed tracheolar networks have been reported but have been shown to be present also in vertebrate tissue (see Wigglesworth, 1931).

PHYSIOLOGICAL ASPECTS OF LUMINESCENCE

Introduction. It was formerly customary for writers on fireflies to include a section or appendix on "physiology." Unfortunately, far too many of these were concerned only with a perfunctory and uncritical exposure of fireflies to various common laboratory reagents. Some of the early work, nevertheless, led to fundamental conclusions in spite of the understandable crudeness of the experiments. In considering modern physiological theories, we shall make extensive use of two of these general findings: the respective influences of the nervous system and of oxygen on luminescence.

The remaining "physiological" literature falls into three classes. The first deals with the actual nature of bioluminescence. These old arguments and experiments over whether fireflies light by means of phosphorus, crystallization, phosphorescence, etc., are interesting historically, but of no relevance to the present discussion since it has been generally agreed for over fifty years that animal light involves an enzymatic oxidation of an organic substrate. The second class of literature concerns the question of whether or not the light is a vital phenomenon, and again this can be dismissed, for no one since Pflüger* (1875) and Bellesme (1880) has doubted that the actual luminescent reaction can occur in the absence of living protoplasm, or even *in vitro*. Finally, there is a huge mass of heterogeneous and unsystematized work which concerns almost every aspect of bioluminescence. Included in this category, for example, are the work on the effects of temperature, the papers on the spectral character of the light (for review, see Buck, 1941, and Grinfeld, 1944), and particularly the reports on the effects of literally hundreds of gases, vapors, acids, bases, salts, poisons, drugs, solvents, excretions, enzymes, etc., on intact fireflies and on isolated organs. All this work suffers from the fundamental ambiguity that it is uncertain whether the agent is acting directly on the actual luminescent reaction or upon a biological mechanism controlling it, as, for example, the nervous

* The facts usually cited as evidence that light-production can be independent of life are the following: (1) Fireflies may still be glowing several days after their apparent deaths. (2) Fireflies which have been quickly dried and kept in absence of air, may glow when moistened, even after several years (*e.g.*, McDermott, 1915). Pflüger (p. 287) countered these two arguments with allusions, respectively, to the sustained irritability of extirpated (surviving) frog muscle, and the viability of cysts of rotifers, etc. Most workers, however, consider these arguments invalid.

system. Furthermore, much of this research is vitiated by failure to specify concentrations, particularly of gases and vapors, and because of the use of single, or only very few, fireflies for each reagent tested. Obviously, we shall be able to utilize only small and isolated fragments from this literature in pursuing the problem outlined in the general introduction. Some of the remaining information, however, is of potential physiological use, though at present one can draw only the most obvious sorts of conclusions, such as that strong protein precipitants extinguish luminescence. The data have further value in showing that one cannot possibly regard the depression or stimulation of luminescence by a given agent as a specific effect, and that one cannot safely argue by analogy, in the sense of saying, for example, that because a capillary dilator increases luminescence, it does so by dilating the tracheoles.

Normal Types of Light-Emission. The difficulties in analyzing the various types of luminescence seen in different kinds of fireflies are increased by the fact that many workers have not distinguished between luminescent behavior in the field and that under laboratory conditions. There are four normal types of light-emission.

The Continuous Glow. This type of luminescence, in which the light is emitted as a glow of rather uniform intensity, usually continuing throughout life, is common in lower organisms such as bacteria and fungi, and is found in the larva of *Phengodes* and in the eggs and pupae of some fireflies. As far as I know, the glow occurs normally in only one kind of mature (adult?) firefly, the larviform female of *Phengodes.* Even here, it fluctuates moderately, following mechanical stimulation.

The Intermittent Glow. In this type of luminescence, the light is emitted as a relatively steady glow which lasts for seconds or minutes. Hasama (1942c) timed the spontaneous activity of a larva of *Luciola cruciata* for 15 consecutive glows and found that the average duration was 20 seconds (range 7 to 60) and the average interval between glows was also 20 (range 4 to 86). Comparable figures were also observed for the larvae of *Pyrocoelia rufa* and *Luciola lateralis* by Hasama (1942b and c), and for *Photuris pennsylvanica* by the writer. In intermittent glowing, the light ordinarily takes several seconds to increase from zero to maximum intensity, and a comparable period to disappear. In larvae of *Photuris pennsylvanica,* however, I have seen the light appear or disappear in approximately a second. Intermittent glowing is apparently under voluntary control, although quite often there seems to be no correlation between luminescence and either activity of the animal or external conditions. Such glows can often be intensified by mechanical stimulation of the animal. Intermittent glowing is characteristic of the genera *Phrixothrix,* *Diphotus,* and probably *Lampyris,*[*] and of the larvae of most fireflies. It also describes well the light-emission in *Pyrophorus,* although Harvey (1931) reported that, after a given glow has passed through its plateau phase and entered the decay phase, a rhythmic fluctuation, with a period of 0.8 to 2 seconds, and an amplitude variation of 25 per cent or more, can sometimes be observed "after the light has nearly subsided." Heinemann (1872) apparently observed the same phenomenon. Harvey ascribed these fluctuations to direct nervous stimulation, although he stated that they might possibly be caused by "some muscular mechanism connected with local distribution of air. . ."

The Pulsation. Hasama (1942a, 1942b) described the light of *Pyrocoelia rufa* and *Luciola lateralis* as being emitted in pulses, averaging, respectively, 6 to 13 and 60 to 110 per minute, with slight sex differences. The pulses are usually fused, thus giving the effect of a continuous light fluctuating fairly regularly. There are also occasional periods of darkness. Schultze's description of a rhythmic ebb and flow in the light of the male of *Lamprorhiza splendidula,* indicates that this species may belong here, too,[†] but no information accurate enough to justify including any other species is available. In regard to frequency, Gerretsen's *Luciola vitticollis* (60-120 per minute) resembles *Luciola lateralis,* as do several of the Oriental and East Indian forms mentioned in many of the extraordinary reports of synchronous flashing (Buck, 1938). However, no further data are available on their types of luminescence.

As will be seen shortly, some forms of flashing also show pulses, but, for the moment, these are considered to be basically different from "the pulsation", because (a) their frequency is very much higher, (b) they occur as part of a rigid, regularly repeated pattern, each burst being followed by several seconds of darkness, and (c) the increase and decrease of luminescence is abrupt. On the other hand, the pulsation may not prove to differ fundamentally from the intermittent glow, in view of Harvey's (1931) photocell string galvanometer studies on *Pyrophorus.* He found that the apparently steady bright plateau glow actually often fluctuates at a frequency of 2½ to 5 cycles per second, though with an amplitude variation (5%) which is too small to be detected by the human eye. For the time being, however, the category is useful for those forms which are neither typical glowers nor typical flashers.

---

[*] It is extraordinary that so few of the early workers troubled to describe directly the type of light-emission of their material, or to differentiate between the emissions of different types of organs, such as the lateral tuberculate and the ventral organs of the female of *Lamprorhiza splendidula.* Spallanzani, and Carus (1864) mention that the larval light of "*Lampyris*" is continuous, but this is opposed by tenuous allusions here and there in the literature. For the emission type of the adult, we have only the indirect statements of Owsjannikow (1868) that the light of the male of *Lampyris noctiluca* "is intermittent, though some lighted for hours"; of Wielowiejski that "after the stopping of intense glow a weak shimmer is seen"; and of Bongardt that "I have never observed that the Lampyridae can suspend their light suddenly". Bellesme states that the female of *L. noctiluca* requires 12 seconds for control, whereas the larva can extinguish its light "suddenly" (2 to 8 seconds). This indicates that at least *Lampyris* is of the intermittently glowing type.

[†] Dahlgren writes of the female of *Lamprorhiza* as glowing continuously and the male as flashing, but it is not apparent that he observed either sex critically.

The Flash. This is the most familiar normal type of light-emission in American fireflies and is seen in many of the common lampyrids (*e.g., Photinus, Photuris, Luciola italica?*[*]). In its simplest form, the flash consists of a burst of light of much greater intensity and much shorter duration than occurs in a glow. Ordinary observations show that the light intensity rises abruptly from zero to a maximum and then declines abruptly again to zero, but the duration of the flash is ordinarily so short that no accurate idea of the phenomenon can be obtained with the naked eye.[†] Using sensitive photocells in combination with amplifiers and a string galvanometer or oscillograph, Brown and King (1931), Snell (1932) and Alexander have carefully studied the "normal" (*i.e.,* laboratory) flashing of *Photinus pyralis* and *Photuris pennsylvanica.* These studies show, among other things, that the duration of the flash is of the order of 0.1 to 0.2 seconds and rather constant, intraspecifically; that the peak intensity and total amount of light emitted per flash are quite variable; and that the augmentation and decay phases are fairly symmetrical (FIGURE 41). The extensive photometric work of Coblentz established 1/400 candle as the light intensity of an average flash of *Photinus pyralis.*

The limitation of most recent experimental work to *Photinus pyralis,* which has a simple single flash, and *Photuris pennsylvanica,* which seems to have a single flash under laboratory conditions, tends to obscure the fact that many much more complex types of flash exist. McDermott (1917) summarized the characteristic flash types given during normal flight by nine American species. Most of these are single flashes of various durations and relative intensities. However, the male of *Pyractomena lucifera* and both sexes of one variety of *Photuris pennsylvanica* emit flashes with multiple peaks, instead of the simple "normal curve" type of luminescence (FIGURE 41). These flashes (better, "coruscations" or "twinkles") can be construed either as due to an optical fusion of separate flashes occurring close together or as rapid fluctuations in peak intensity of a single flash. I have records similar to McDermott's of the flashing characteristics of about forty species of Jamaican fireflies, of which eight show a coruscating type of flash. These multiple flashes differ interspecifically in duration, and in the frequency of the individual peaks, and in some species occur in complex, but constant, combinations with single flashes. The maximum frequency of the oscillations in intensity during a coruscation is not known, but is certainly not much less than the critical flicker fusion frequency of the human retina (probably from 25-30 cycles per second under the conditions of observation). Lund's description of

---

[*] In *Luciola italica,* Emery described the "increase and decrease of the light in short regular intervals," and Verworn (1892) described the "rhythmic intermittency" of the light of both sexes in flight. The frequency was 60 to 80 per minute according to Verworn, and 80 to 100 per minute according to Peters (1841). Verworn mentioned a weak continuous glow, given while resting, and stated also that the light was not completely extinguished between flashes. This description fits pulsing as well as flashing, but Geipel's "strict rhythm of flashing and extinction" indicates that *L. italica* is a flashing type.

[†] Occasional individuals of flashing species show a very faint glow which persists between flashes.

"regular, rapid and numerous changes in intensity" in the flashing of two Jamaican species of *Photinus* almost certainly refers to such coruscations.

It is possible that even the apparently single flashes of some species may show finer cyclic oscillations similar to those discovered in *Pyrophorus* by Harvey (1931). In *Photinus xanthophotis catherinae,* which has a very large organ and a flash lasting about half a second, I occasionally saw, with peripheral vision, very slight and very rapid fluctuations. However, it is also possible that a subjective effect is involved in this, since it is well known that, with light of very low brightness, the eye tends to shift its retinal fixation point rapidly. Such an effect might also be the explanation of the *"ébranlement particulier"* of the "retina" which Bellesme reported in close observation of the light of *Lampyris* in a darkroom.

In many flashing types of fireflies, light-emission is used in astonishing systems of mating signals in which the male flies about, flashing at regular intervals, while the female, usually at rest, flashes in response to his signals (Osten-Sacken, 1861; McDermott, 1910, 1911a, 1912, 1917; Mast, 1912; Buck, 1937b). These signal systems are remarkably precise in their time relations and differ characteristically in different species. In normal flight, the male produces his luminous unit (whether single flash, coruscation, or complex) at regular intervals, and this emission pattern is so characteristic that the males of each of the dozen or more species which may be active at night in the field at the same time can be identified reliably from observation of their flashing, alone. As might be expected, the frequency of flashing rises with temperature (Snyder and Snyder, 1920; Buck, 1937b).

In summary, these four types of light-emission can be interpreted as the results of four progressively more effective modes of control. The continuous glow is an indication of the inability of the animal to prevent the luminescent reaction, or even to change its rate materially. In the intermittent glow type, the organism can interrupt photogeny, but only slowly, and the plateau level of glow presumably represents the condition with no control operating. In the pulsation, a further refinement of control may be assumed to have developed, so that now the light can be made to fluctuate fairly rapidly. In the flash, finally, the control mechanism achieves its highest development, as indicated by its ability to bring about bursts of light of very quick accretion and decay, of short duration, and with complete extinction between even closely spaced flashes. Since, in most flashes, nothing resembling a "plateau" is reached, it is uncertain whether or not the peak of the flash represents maximum possible luminescence.

With the above hypotheses in mind, it is interesting to review these various types of light-emission in connection with the morphological studies already discussed. In the continuously glowing *Phengodes,* the luminous cells lack any specialized tracheal supply and presumably obtain oxygen from the body fluid with which they are bathed. Tracheae are also absent in eggs, and presumably also in pupae, because of histolysis. In the organisms with intermittent glows, we find, exclusively, organs of Types 2 and 3, with a branched and tapering tracheal supply lacking end-cells. In the fireflies with the pulsating type of light-production, we find end-cells present, but with arborescent tracheal branching, as described for organs of Type 5. Finally, in the flashing type, we find the cylinder, the highest morphological development in the sense of precision and complexity of organization. Thus, there seems to be a striking correlation between the type of tracheal supply in the photogenic organ and the normal type of luminescence. In particular, the presence of end-cells seems to be associated with the ability to produce a flash or pulse. This has been pointed out previously by various workers on the Type 6 organ* and will be discussed further in connection with end-cell physiology. Whether or not the proposed distinction between the pulsation and the "true" flash will prove valid must await careful observation on further species having the Type 5 tracheal arrangement. At present, I am not able to suggest any very convincing anatomical reason why the Type 5 organ should be less abie than the Type 6 to control luminescence abruptly. Control might be connected with the cylinder itself, but it might also be related to some other anatomical feature such as, for example, the differentiated zone of the photogenic cytoplasm, or tracheolar anastomosis, neither of which seems to have been recognized in firefly photogenic organs of Type 5.

Localization of Luminescence. Microscopic observation of the active photogenic organ, preferably in the living animal, provides the only acceptable evidence of where light-production is localized. A number of investigators have reported work of this sort. Except for Weitlaner (1909), no one, since Kölliker first demonstrated the two-layered structure of the organ, has doubted that the light is produced in the "photogenic" layer, although certain workers (*e.g.*, Wielowiejski) thought that the reflector layer, too, was slightly luminescent. Almost certainly, this idea was fostered by contamination of the reflector layer with photogenic material or by diffusion of light from the underlying photogenic layer.

Spallanzani was probably the first to observe the surface of a glowing light organ. He reported having seen many tiny points of light. Kölliker and Schultze observed the same "minute sparks" in *Lamprorhiza,* and Schultze, in addition, believed that the points corresponded to his newly demonstrated end-cells. Schultze is generally credited with the claim that the light is confined to the end-cells, but actually he referred to the light as "beginning" in the end-cells. Since he also suspected a connection between nerves and "parenchyma" (photogenic) cells, Wielowiejski and Bongardt have argued that this implies his believing the photogenic cells to luminesce in addition. Bongardt and Lund described just such a spread of light from minute flashing points into a general glowing of the whole organ, and stated that the positions of the individual spots were constant, though their intensities were not. Bongardt also claimed that the number of end-cells, as determined from the points of reduction of osmic acid, far exceeds the number of lighting points, indicating that the end-cells are not the luminous points. This argument is not necessarily relevant, since not all the end-cells would need to be active at once.

---

* This is apparently what Hasama (1942c) was referring to in his statement that "many authors have ascribed the strongly rhythmic blinking of tropical fireflies to the concentric arrangement of the end-cells." (Translation.)

The localization of luminescence has been much more extensively studied in the Type 6 organ than in *Lamprorhiza* and *Lampyris*. Emery set the general pattern for all later workers when he described the surface of a glowing (not sparkling) *Luciola italica* organ as showing a pattern of luminous rings, the dark centers corresponding to the cylinders. The same sort of structure has been observed in *Photinus marginellus* (Townsend), *Photuris pennsylvanica* and three Jamaican Photini (Lund), and in *Photinus pyralis* and *Photuris pennsylvanica* (Alexander). Emery discovered, in addition, that the luminous rings were not uniform but showed spots which lit up and went out irregularly and were, he thought, constant in position. Emery is a little vague about the exact localization of these spots, but it appears that he believed the light to occur at the contact between the end-cells and the photogenic cells. Lund confirmed Emery's observations in all respects and added the finding that the end-cells did not luminesce. He also discovered that when an organ was glowing brightly enough so that the individual rings fused and made the whole intercylinder area luminescent, the brightest region was in the differential zone at the periphery of the photogenic cells. Alexander, in turn, confirmed Lund's report on the localization of the bright points of light at the bifurcations of the tracheoles, and on the fact that they fire repetitively and asynchronously. This rapid, irregular, scintillating or "spinthariscope" type of local luminescence has also been observed by Wood (1939) in a firefly poisoned by a spider,* and by Kastle and McDermott (1910) and Alexander following a variety of treatments, of which strychnine injection was most effective. In addition, I have seen it following distilled water injections and cyanide vapor. Besides the glowing rings and sparkling points, "phosphorescent clouds" which sweep across the organ in waves have been described (Emery, Kastle and McDermott, Lund). According to Lund, they originate deep in the tissue.

The observations just described were used by Lund to support his claim concerning the enzymatic activity of the periphery of the tracheole, which has already been mentioned, and by Emery to further his idea that the photogenic material "secreted" in the light cells is transported to and burned at or in the end-cells. It should be kept firmly in mind, however, that no details of any kind can be made out in an organ which is flashing, or even exhibiting a bright glow, and that the conditions under which the luminous rings, scintillations, and cloudy waves are seen are distinctly abnormal. Moreover, it must be remembered that none of the findings have any necessary relevance for organs of Types 2 and 3, where end-cells and cylinders are lacking.

---

* Steche (1908) observed normal intermittent flashing in a firefly apparently paralyzed by spider bite.

In contrast to light organs which contain end-cells, the organs of Types 2 and 3 show only a uniform and structureless glow. Dahlgren and others have pointed this out as characteristic of *Pyrophorus,* "the larvae of all Lampyrids", "the females of *Lamprorhiza splendidula* and all species of *Phengodes*". From personal observation, I can confirm Dahlgren's statement on *Pyrophorus* and the larva of *Photuris pennsylvanica.* I would expect the statement about *Lamprorhiza splendidula* to apply only to the lateral tuberculate organs. In *Phengodes,* I have described the light as coming from minute separate spots corresponding in number, size, and position to the huge oenocyte-like cells seen in histological preparations (FIGURE 15; Buck, 1946a).

The theories of Schultze and of Emery, which gave the end-cell a primary role in photogeny, were based on two observations: the presence of points of light in the general region of the end-cells in the living light organ, and the affinity of the end-cells for osmic acid vapor. They also involved the assumption that a high reducing power for osmium indicates a high affinity for oxygen. We have already seen that Lund denied that the end-cells actually luminesce, and that several workers have described the spread of luminescence throughout the photogenic cytoplasm. Nevertheless, observations on the living organ are not precise enough to rule out the possibility that the end-cells, or other inter-cylinder material, may light in addition to the photogenic cells. This is particularly relevant since so many reports speak of the luminescence beginning at the edge of the cylinder and since Lund himself admitted that the end-cells are often situated in little bays hollowed into the photogenic cytoplasm. The point about the site of reduction of osmium, which Emery regards as the "*experimentum crucis,*" and which Lund also uses to bolster up his theory of the enzymatic activity of the tracheolar wall, has been dealt with on page 409. The particularly heavy deposit does not necessarily prove that there is anything unique about the end-cell, but merely that it stands at the first spot permeable enough to allow the vapor to escape. The slight penetration into the tracheoles would then be caused by the osmium being reduced as fast as it arrives at their proximal ends. Wielowiejski has argued, in addition, that a high affinity for oxygen does not necessarily identify the site of photogeny, and has made the interesting suggestion that the end-cell acts in a manner analogous to the red blood cell and "stores" oxygen. Wielowiejski's and Emery's theories thus each provide that material, freed in the photogenic cells by nerve stimulation (the control), moves toward the end-cell, by a method not elaborated, and luminesces near it because the highest concentration of oxygen is found there. However, no evidence exists to show that any oxygen carrier is present in or around the end-cell, and in any event its activity would be an effect, rather than a cause, of luminescence.

Tracheolar Properties of Interest in Regard to Control of Luminescence. Creighton injected adrenalin into *Photuris pennsylvanica* and observed a bright glow. This, as we have seen, is the usual response to dozens of agents. On the basis of histological studies, he claimed that the adrenalin effect was not hormonal but was due to contraction of muscle fibers in the tracheal end-cell which enlarged the part of the tracheole within the end-cell, as postulated by Dahlgren. However, according to a personal communication cited by Alexander, it appears that Creighton's preparations showed a dilation not of the end-cell lumen but of the whole tracheole. We have, then, the possibility that luminescence may be controlled by an active or passive dilation and constriction of the tracheole, such as Gerretsen suggested as an alternative to end-cell activity.

Many have regarded the fact that the tracheoles do not dissolve in caustic solutions as proof that they are composed of chitin, although this has not been confirmed by more specific tests. If the walls of the tracheoles were composed purely of chitin, they would be unlikely to be very dilatable. However, as Richards (1947) has pointed out, protein often bulks large in cuticle and could satisfy most conditions for extensibility. It would be desirable to have better evidence of the chemical composition of the tracheoles, but color tests would not be conclusive, in my opinion, because of the difficulty in getting a strong enough color in a structure as thin as the tracheolar membrane. For many years, the absence of visible spiral thickenings in the walls of the tracheoles was regarded as strong evidence that they differ from the ordinary larger tracheae. However, Richards and Anderson (1942) have shown by electron microscopy that honeybee tracheoles have spiral thickenings at least down to a diameter of 0.2 micron. The presence of thickenings in a tracheole might reduce the extensibility of the wall, but would not prevent tracheol collapse, as Richards and Anderson have shown.* However, if Creighton's finding of tracheoles dilated by adrenalin is accepted, it indicates that the chitin content is low.

There is an opinion that tracheoles must differ in composition from tracheae because they are permeable to gases and liquids. This idea probably stems from Krogh's (1919) evidence that the diffusion rate of oxygen through chitin is only about one-thirtieth of that through water, and about one-tenth of that through tissue. However, if it is recalled that the membrane of the tracheole is only between 0.005 and 0.01 micron thick (Richards and Anderson), it will be seen that simple transfer could go on readily even through chitin. Wigglesworth (1930) and others have, in fact, shown that tracheae are freely permeable to oxygen and nitrogen.

---

* **References to additional literature on tracheoles will be found in Wigglesworth (1931 and 1939), and Richards & Anderson.**

Hoskins (1940) has called attention to one possibly valid distinction between tracheae and tracheoles, which is that the former are as a rule distinctly hydrophobic, whereas the latter are often strongly hydrophilic. Even here, some confusion exists, since it has been reported that the tracheae are hydrophilic in newly moulted insects (Wigglesworth, 1938b) and, conversely, that oil which enters the tracheae eventually penetrates into the tracheoles (Hoskins).

It is almost universally held, either explicitly or implicitly, that the principal function of the tracheoles in the light organ is to conduct oxygen to the photogenic cells. It should be emphasized that, however reasonable this view may appear, it is only an assumption. Moreover, there has not even been a serious attempt to define, theoretically, the conditions under which oxygen conduction might take place. Among the more obvious items of needed information are data on whether transport occurs by diffusion or by flow, and whether the tracheoles contain air or liquid. A further discussion of theoretical aspects of these questions will be found in the section on end-cell physiology. For the present, we can conclude only that, if transport is by diffusion, the tracheole loses most of its meaning if it is assumed to be filled with liquid, because the diffusion of oxygen down a column of water is no faster than it would be in free water, and only about three times as fast as through dense connective tissue (Krogh). Only by containing air, in which oxygen diffuses thousands of times as fast as through water, could a tube confer any advantage in supplying oxygen. To be sure, if the tracheole were absolutely impermeable to oxygen for some of its length, access of oxygen to parts of the photogenic cytoplasm could be prevented or delayed even if the tracheole were liquid-filled, but it is hard to see what advantage would be conferred thereby and, moreover, we have already seen that impermeability to oxygen is extremely unlikely.

Another possibly significant point is Wigglesworth's (1930) observation that air bubbles, trapped between two columns of oil advancing toward each other in a trachea, are completely squeezed into solution, in the hemolymph or tissue surrounding the trachea, by the surface tension of the oil. I have seen the same thing happen in dipteran larvae immersed in kerosene. An astonishing feature of the process is its speed. The factors involved are too complex to permit one to say off-hand whether the same phenomenon would occur in tubules of the dimensions and structure of tracheoles, or, if so, whether it would be rapid enough to meet the requirements of photogenic control. However, the principle should be kept in mind as a possible method of forcing oxygen into cells. If, for example, the tracheal end-cells were to "close", and if water moved up the tracheole proximally, it might drive some of the contained air into the photogenic tissue. This would be the reverse of the theory of osmotic control of luminescence to be discussed in a later section.

The Nervous System and Luminescence. A connection between the nervous system and luminescence has long been accepted by most investigators, on the basis of the anatomical evidence already discussed and because of a variety of experimental work. It was shown very early that decapitation, or cord section anterior to the organ, results in immediate cessation of voluntary flashing or glowing, though a dim constant luminescence may persist, or reappear some time later (Macaire, 1821; Peters, 1841; Verworn, 1892; Dubois; Prowazek, 1908; Lund, Williams). Likewise, luminescence is initiated or increased by stimulating the nerve cord[*] mechanically (*e.g.*, Heller, 1853; Verworn), or electrically (Macartney, 1810; Macaire; Todd, 1826; Joseph, 1854; Kölliker; Owsjannikow, 1868; Bellesme; Heinemann, 1886; Dubois; Fuchs, 1891; Steinach, 1908; Lund, Gerretsen, Perkins, 1931; Snell, 1932; Brown and King, Alexander), even after section proximal to the point of stimulation.[†] Direct electrical stimulation of the photogenic tissue likewise causes lighting. Nervous influence is also suggested by numerous observations that anesthesia suppresses voluntary control of luminescence,[‡] and by the apparently stimulating effects of spider venom (Wood), various neurotoxins (Kuhnt), and DDT (page 75). In all these instances, however, direct action of the agent on the photogenic tissue, or on the tracheal system, cannot be excluded.

In contrast to the above work, Owsjannikow observed no diminution of light in isolated organs after 1½ hours' soaking in curare and strychnine, and concluded that the nervous system is not involved in luminescence. Bongardt repeated this experiment, using the entire animal, and reported no effect in 12 hours. Aside from the questions of penetration and of whether vertebrate poisons would be expected to act on insects (raised also by Bongardt), this conclusion involves a type of *non-sequitur* encountered frequently in the older work, and well illustrated also by Bongardt's argument that fireflies cannot stop their light suddenly, because dead ones continue to glow for 10 to 20 days. Owsjannikow and Bongardt thus failed to see that what happens in the injured or dead animal may be entirely irrelevant to the question of whether the nervous system, or anything else, permits voluntary control of luminescence in the normal living animal.

---

[*] Because of the minute dimensions involved, it is extremely likely that simultaneous stimulation of other tissues near the operated region was not excluded, particularly in the earlier work.

[†] Some of these papers report considerably more than the bare observation. They are interesting from the standpoint of comparative electrophysiology.

[‡] However, luminescence itself may persist for long periods under conditions where voluntary neuro-muscular activity is suppressed or abolished entirely.

Harvey's (1931) photocell-string galvanometer records on the glowing of *Pyrophorus,* which have already been discussed, indicate that the light is under nervous control, particularly as Harvey was careful to exclude respiratory and pulse movements. A better test could be made by recording simultaneously the action potential pattern of the ventral nerve cord, and the luminescence, particularly if done on several species with differing flash patterns. It may be of interest here to mention some other potentials associated with luminescence, though not necessarily with the nervous system. Many years ago, Kölliker obtained some "not quite constant" evidence that lighting fireflies deflect a "multiplier" (galvanometer) more than do non-luminescent ones. In view of Kölliker's primitive apparatus, his report does not call for extensive consideration. Recently, however, Hasama has reported that, in a number of Korean species, the light organ itself produces a monophasic action potential during activity. The luminous segments are electronegative to the non-luminous. In the larva, the potential pattern is continuous, and synchronous with light emission. In the adult, it exhibits rhythmic or cyclic fluctations, the frequency of which corresponds satisfactorily with that of the pulsing of the light in *Pyrocoelia rufa* (13 per minute), but poorly in *Luciola lateralis* (48 per minute for the potential, 60 per minute minimum for the light). With the electrodes on two non-luminous segments, instead of on one luminous and one non-luminous, no potential difference was detected, thus indicating that muscle potentials, etc., are not involved.

A number of further lines of evidence indicate indirectly that the nervous system is concerned with the control of luminescence. For example, the rigid characteristic species flash-patterns already referred to, as well as the inherent diurnal rhythm of luminescent activity* (Allard, 1931; Perkins; Rau, 1932; Buck, 1937a), point to the existence of at least an involuntary center of nervous control. Indeed, Verworn, some sixty years ago, postulated in detail that normal luminescence is under absolute control of an automatic nervous center in one of the two most anterior ganglia of the cord. Though similar hypotheses have since been used by several writers in attempting to explain synchronous flashing (see Buck, 1938), work on the mating signals shows that a high degree of voluntary control can also be exercised. Briefly, these signal systems enable the female of one species, or an artificial light operated in a specific way, to attract males of the same species without attracting males of another species. Conversely, males of a given species signal only to a female of the same species (or to a properly operated artificial light) and ignore other males of their own species, both males and females of other species, and all improperly executed artificial signals. In *Photinus pyralis,*

---

* This periodicity may play an important and unsuspected part in certain experiments, since Maloeuf (1938) and Alexander have claimed that it influences reactions as basic as the appearance of luminescence in high partial pressures of oxygen.

I have shown that the ability of the male to recognize the female depends on the fact that she always replies to his flash after a particular time interval (Buck, 1937b).

The fact of nervous influence, however, does not solve the problem of the control of luminescence, since the effect could either be direct (*i.e.*, stimulation of the photogenic cells), or indirect (*e.g.*, stimulation of the end-cells; oxygen regulation). On this question there is no dearth of opinions, pro and con, but valid empirical evidence has proved to be extraordinarily hard to obtain.

Evidence that oxygen regulation is the primary factor in control of luminescence is, of course, opposed to the idea of direct nervous stimulation of the photogenic cells. This evidence will be analyzed later in detail (page 441), but it may be said here that it is insufficient to exclude the possibility of direct nervous control.

Heinemann (1886) reported that electrical stimulation of nerves anterior to the abdominal photogenic organ of *Pyrophorus,* after all other tissues of the trunk had been transected, did not induce luminescence. From this he concluded that the nerves act by way of some intermediate tissue. However, since it was a negative type of experiment, and since Lund, in repeating it, obtained precisely the opposite result, it cannot be regarded as decisive. It is also difficult to reconcile with numerous reports of luminescence obtained by electrical stimulation of isolated abdomens of lampyrids, and with the experiments of Fuchs on localized stimulation of the photogenic organs and nerves of *Pyrophorus.*

Considerable evidence has been adduced in favor of direct nerve action. Lund's transection experiment has already been mentioned. However, Lund's acceptance, on the basis of this and other work, of "primary control of the organ . . . by nerves in direct connection with the photogenic tissue" is greatly weakened by a number of apparent contradictions in his presentation. For one thing, it appears that, in spite of denying oxygen a main role in control, he actually considered the end-cell to be of prime importance. For another, he emphasized the well-known fact that the photogenic tissue itself "is irritable and responds locally to mechanical stimuli", overlooking the probability that such disturbances also facilitate access of oxygen. Steinach, using induction shocks applied directly to the photogenic organs of decapitated specimens of *Lampyris,* claimed to have demonstrated summation of subliminal stimuli for luminescence. He attributed the summation to a direct effect of the nerves on synthesis of substrate for luminescence, and rejected the alternative idea of simple neural summation, although his reasons for so doing are unconvincing. Bellesme also supported direct nervous stimulation of synthesis on the basis of an experiment in which he found that fireflies cut open in air luminesced sooner, if they had first been "stimulated" electrically in the absence of oxygen, than if they had not. However, aside from the unavoidable variations in manipulation, the temporal variability in responses of fireflies under such conditions is ordinarily so large that a conclusion cannot be accepted until verified on a considerable number of specimens. Moreover, no allowance was made for the probable effects of the anoxia upon the nerves. Additional details concerning the work of Steinach and Bellesme will be found on pages 438 and 441. Work of Snell bearing on the same problem is considered on pages 442 and 443.

Perhaps the strongest evidence of direct nervous control of luminescence is furnished by the numerous experiments on electrical stimulation of *Pyrophorus* and of firefly larvae. Here, the finding of increased intensity of luminescence with increased intensity of stimulus is less easy to attribute to oxygen control, since these forms lack end-cells.

The problem of direct *versus* indirect nervous control of luminescence is thus unsolved, although there is much evidence in favor of the former alternative. Unfortunately, the effects of oxygen and nerve action seem inseparable, experimentally. Under abnormal laboratory conditions, it is clear that nerves are not essential for luminescence, since fireflies will glow under a variety of conditions incompatible with nerve action, or even with life, if oxygen is present. Nevertheless, as we shall see, there is no conclusive evidence that oxygen is ever limiting in normal flashing, and there are many indications that control is much more circuitous than by direct regulation of oxygen access to the photogenic tissue. A number of hypothetical mechanisms of intracellular control, such as might be set off by direct nervous stimulation, are discussed on pages 434 to 436. Moreover, it should be kept in mind that the fact that the experiments discussed are inconclusive, does not disprove direct nerve action. The interpretation of Bellesme and Steinach may yet turn out to be correct, even though it does not necessarily follow from their experiments.

THE CONTROL OF LUMINESCENCE

Historical. Before entering on the detailed discussion of modern work and theories on control of luminescence, it may be interesting to consider, briefly, a few of the older theories proposed during the long history of the study of fireflies, and now abandoned. Perhaps the oldest and most persistent idea was that luminescence is controlled in some way by "respiration", by which apparently was meant that the muscular respiratory movements pumped air into the photogenic organ (Joseph, 1854; Faraday, 1814; Carrara, 1836; Siebold, 1848; Leydig, 1857; Seaman; Heinemann, 1886; Watasé). Several of these observers maintained that the breathing movements are synchronous with light-emission. However, breathing movements could hardly be the only control, unless the light pulsated day and night or the insect stopped breathing by day, neither of which alternatives seems likely. Elaborations on the idea were proposed by Carrara, who claimed to have found an air tube leading from the mouth to an abdominal bellows or vesicle; by Heinemann (1886), who envisaged the abdominal musculature as pumping air through the "hiatuses" in the photogenic layer of *Pyrophorus;* and by Seaman, who suggested that the spiracles were first closed and air was then forced into the organ by abdominal contraction. A serious stumbling block for these theories is the fact that most modern workers have been unable to observe any visible muscular movements coinciding with luminescence (Lund: Harvey, 1931; Hasama, 1942a). Likewise, Heinemann's theory has little support, since Geipel has shown that the musculature of *Pyrophorus* is in no way different from that of non-luminous elaterids, and since, as previously stated, the existence of the spaces in the photogenic layer is very dubious. The same objections apply to Dubois' theory, according to which not air but blood was supposed to be percolated through the photogenic organ of *Pyrophorus.* The circulatory system was also implicated by Carus, who maintained that the rhythm of luminescence corresponded to the pulse, and by Tozzetti, who suggested that the tracheoles carried part of the blood.

Another theory was to the effect that the light-emission was controlled by withdrawing the light organ into the interior of the abdomen, where it was concealed by other viscera (Carradori, 1797; Müller, 1805; Owsjannikow, 1864).

Deductions from Kinetics Analysis. By high-speed recording, Brown and King, Snell and Alexander obtained time-intensity curves of the flashing of *Photuris pennsylvanica* and *Photinus pyralis.* Snell pointed out that the form of these curves is determined by three variables: the number of photogenic units active, the degree of activity of each unit, and the degree to which each unit is stimulated.

If these curves represented the response of a single photogenic unit, or the simultaneous responses of all the photogenic units in the light-organ, their analysis could reveal valuable information on the luminescent reaction and its control mechanism. From the form of the accretion and decay phases, it might be possible to deduce, for example, the number of reactants involved in the rate-determining reaction and the number of units active; and to ascertain whether diffusion or "phase-boundary removal" (see page 436) brings the reactants together; whether or not the unit response is of the all-or-none type; whether the control mechanism "opens" by relaxation or contraction; etc. Brown and King, in fact, suggested a relation between the supposed logarithmic form of the decay phase of the flashing curve, and the similarly logarithmic form of the decay phase of the luciferin-luciferase reaction (*Cypridina*) *in vitro.* According to Snell, however, a logarithmic form is the exception rather than the rule. Snell found the durations of normal flashes very constant, the intensities highly variable, and even the reaction velocities (slope of decay phase) different.

The two segments of the luminous organ of the male of *Photinus pyralis* contain about 6,000 cylinders, each of which encloses a minimum of 80 to 100 end-cells, which, in turn, give off twice that many tracheoles. The tracheoles supply an estimated total of 15,000 photogenic cells, each of which is in contact with two cylinders (FIGURES 13 and 31). In studying the glows produced by various agents, I have noticed many times that single cylinders may light as individual units—that is to say, the surface of the organ may show minute isolated "doughnuts" of light which glow on and off independently. We must allow, then, for at least 6,000 photogenic units. If we accept Lund's and Alexander's claims that the end-cell (region) can flash independently of the photogenic cells, as in "scintillation." the number must be increased to something like 600,000. In either case. there is a strong likelihood that the time-intensity curve for the flash of the whole organ represents not the result of an absolutely synchronous firing of all the units, but the statistical result of the firing of units slightly out of phase. This is supported by the suggestive resemblance of the time-intensity curve to the normal-distribution curve. If the curve is indeed a statistical one, we can draw the very important conclusion that the duration of the luminescence produced by a single photogenic unit must be far less than the 0.15 second average for the mass flash. We have no way of knowing what the duration of the flash of the individual luminescent unit (as distinguished from its glow) may be, but it can hardly be more than one-tenth of that of the collective flash. The importance of this inference is that it imposes much more severe conditions upon the control mechanism. We now require a mechanism capable of producing isolated flashes of light with durations of the order of 0.01 second, or less. The same conclusion is indicated by the observations previously reported on the repetitive or oscillating types of flash seen in some Jamaican fireflies and in *Photuris pennsylvanica* in the field, although, here, there remains the possibility that the coruscations could be produced by groups of units firing successively in relays.

Mechanisms of Intracellular Control. We have seen that there is no agreement on whether flashing is controlled primarily by the nervous system or by the oxygen supply system. There is also no agreement on whether these two systems operate directly, *i.e.*, on the photogenic cell, or indirectly, *i.e.*, on each other. It is therefore appropriate to consider what general types of mechanism within the photogenic cell might be available as a basis for assumed "direct" effects of nervous or oxygen stimulation.

The protozoan *Noctiluca* is described (Quatrefages, 1850; review in Harvey, 1940) as emitting a bright flash, consisting of the momentary lighting of a multitude of tiny sparks scattered through the protoplasm, each representing a granule of photogenic material. The flash is not repetitive or voluntary, in the sense of being initiated internally, but is a direct response to some environmental stimulus (osmotic, thermal, chemical, mechanical, etc.). The synchronous lighting of the cytoplasmic granules during the flash is presumably attributable to the fact that the stimulus reaches the cell through the ambient water and, hence, almost simultaneously at all points on the surface. The bare essentials for controlled luminescence thus seem to be, first, the presence of a chemiluminescent system in the cytoplasm; second, a source of the necessary raw materials (here obtained from the surrounding sea water); and, third, a stimulus which will set off the intracellular reaction.

Harvey (1940) has pointed to the fact that a single cell like *Noctiluca* can flash as showing that complicated physiological mechanisms, such as those in the firefly, are not necessary for the control of luminescence. He thus regards the response of *Noctiluca* as quite similar to the responses of other cells (*e.g.*, muscle) to direct stimulation, and implies that the response of an individual photogenic cell of the firefly need not, intrinsically, be any more complex. According to Harvey's view, therefore, the nervous system would correspond to the environmental changes which stimulate *Noctiluca,* and, in the firefly, would represent the anatomical and physiological answer to the problem of controlling the responses of thousands of *Noctiluca*-like photogenic units. The tracheal system, likewise, would stand in the same relation to the photogenic tissue as the sea-water to *Noctiluca,* and would be an arrangement for attaining, throughout a massive tissue deep within a body, conditions of aqueous diffusion of oxygen comparable to those of a free cell suspended in aerated sea-water. We may, therefore, think of *Noctiluca* and the photogenic cell of the firefly as presumably equivalent in regard to their intracellular control mechanisms.

Before considering possible intracellular controls, however, it is well to emphasize that the nervous and tracheal systems of the firefly, subsidiary though they may be, have important influences on luminescence. They are, therefore, as much a part of the control problem as the intracellular activities, and can be investigated independently of the latter.

A number of intracellular control mechanisms have been proposed, all on largely speculative bases. Perhaps the simplest of these, which may be called the "reactant limitation hypothesis", postulates that luminescence is controlled by limiting one of the four fundamental components of the luminescent reaction. With respect to oxygen, this hypothesis is equivalent to direct external regulation, but in regard to the other three essential reactants, even our very limited knowledge of the chemistry and enzymology of luminescence is sufficient to suggest various intracellular processes which might limit light-emission. Various reports indicating control by substrate synthesis, or storage and release, are considered on pages 437 to 440.

A second intracellular control mechanism has been proposed by Harvey on several occasions, and may be called the phase-boundary hypothesis. This hypothesis provides that all the essential reactants are present in the cell at the same time, but are prevented from reacting with each other by phase-boundaries. When these boundaries break down, as might occur as a result of nervous stimulation of the cell, the reactants can come together and produce light. Phase-boundary phenomena are well known in physical chemistry, and there are a few suggestive physiological and biochemical analogues, but we are still far from having a clear idea of how potential reactants are segregated from each other in the living cell. In fact, phase-boundary changes may only be some subtle form of reactant limitation. A single test of the phase-boundary hypothesis has been made on fireflies, but, as described on pages 442 and 443, the results were not conclusive. The similarly inconclusive experiments of Bellesme and Steinach on direct nerve action (pages 432 and 437) can also be interpreted as consistent with phase-boundary control. For the present, however, the concept must be considered simply as a theory, albeit a reasonable and stimulating one.

A third hypothetical intracellular control mechanism is one which makes use of a possible competition for oxygen between respiration and luminescence. This hypothesis, which has a number of variants, and which has not been tested experimentally, will be discussed on page 448.

None of the above discussion of intracellular mechanisms bears on the validity of the commonly held view that nervous control is direct and the tracheal system merely a means for ensuring that adequate oxygen is always available for luminescence. We shall be able to judge this better after having considered the evidence on oxygen limitation.

In conclusion, on *a priori* grounds it appears that intracellular control mechanisms would have a considerable advantage over external control mechanisms (*e.g.*, tracheoles or end-cells) in regard to speed of action. This might be crucial in the control of high-frequency light-emissions, such as are seen in the coruscating types of flashing. On the other hand, by making the control of luminous intensity more or less independent of changes in external oxygen tension, doubt is cast on the generally accepted significance of certain anatomical features of the photogenic organ, particularly the end-cell.

Water as a Possible Limiting Factor in the Control of Luminescence. Some of the earliest experiments demonstrated that if a firefly or an extirpated organ is quickly and thoroughly dried, it can be preserved in a non-luminescent state for long periods and will glow again when moistened (Spallanzani, Carradori, Macaire, Kölliker, Carus; Owsjannikow, 1868; Dubois, Bongardt, Kastle and McDermott; Hasama, 1942a). However, it seems extremely unlikely that water lack is ever a direct factor in controlling light-emission in the living animal, since this would appear to demand a high-frequency cyclic dehydration and rehydration of the photogenic cells, and since luminescence continues for a considerable time in minced organs, where dehydration must be severe. This, then, is a good example of a method of limiting the luminescent reaction *in vitro* (or at least in the absence of living protoplasm) which is not involved in the normal control of photogeny.

Luciferin and Luciferase as Possible Limiting Factors in the Control of Luminescence. In discussing various theories of photogenic control by limitation of reactants, it should be understood that the terms, *photogenic material,* or *substrate,* could refer either to luciferin or luciferase, since there is no evidence which points specifically to either.

If either luciferin or luciferase were not in the luminous organ, the problem of control could be simplified. It was, in fact, a convenient feature of Dubois' control theory that the luciferase was brought to the organ by the blood which circulated through the interstices of the photogenic layer and there met the luciferin. In all modern speculations on luminescence control, however, it has tacitly been assumed that these two reactants are formed in the photogenic layer and remain there, and that control is therefore a cellular problem. Moreover, most writers have assumed that photogenic materials are usually stored in the photogenic organ in excess, and that synthesis of substrate has no role in rapid control of luminescence. This view is based on the aforementioned observations that dead, minced, or even dried and remoistened fireflies or organs will glow for many hours. Since, however, the enzyme luciferase seems to survive even drying, there is no *a priori* reason why synthetic or other enzymatic methods of releasing substrate might not also survive.

In considering luciferin or luciferase as a possible limiting factor in luminescence, it is necessary to distinguish clearly, as Bellesme and Gerretsen have done, between the reactions which form the luminous material and those in which luminescence takes place. This distinction has often been overlooked. It is also necessary to differentiate carefully between mechanisms which operate to keep luciferase or luciferin separated (both being present) and those which limit the rate of formation of one or the other. The phase-boundary control mechanism, based on the former hypothesis, has already been discussed (p. 432). The latter hypothesis, according to which the rate of synthesis of "photogenic material" in the photogenic cells may become limiting, has also been debated, though not always in a clearly defined form. Steinach made implicit use of the hypothesis when he reported that stronger electrical stimulation produces a stronger light because more "light-stuff" is formed. Bellesme supported the same hypothesis explicitly when he postulated that the photogenic cells secrete a gaseous product (phosphine!) which lights as soon as it comes in contact with air. On the basis of the rather dubious claim that a thoroughly crushed glowworm does not glow (when it ought to, since its interior is exposed to air), Bellesme postulated, in addition, that no lighting material is stored in reserve. In further support of this, he reported that there was a latent period between stimulus of the animal and the beginning of luminescence, and that this decreased if the firefly was first stimulated electrically in the absence of oxygen, so as to build up a store of luminous material. Delay might very well be seen in glowing forms, where the build-up of luminescence is slow, but, as we shall see, it is more likely to have another explanation. Additional details of, and objections to, the experiments of Steinach and Bellesme have been presented on page 432.

The inference that storage is shown by the fact that a firefly induced to flash for prolonged periods eventually becomes unable to luminesce, is too fallacious for extended consideration.

The question of storage of photogenic material was also considered by Wielowiejski and Bongardt, though in a very circuitous fashion. Wielowiejski opposed the idea of storage of substrate on the grounds that otherwise it would be impossible to explain how fireflies are able to extinguish their light voluntarily. This idea was based on Wielowiejski's convictions that neither end-cells nor oxygen are directly concerned with the control of luminescence, and that it is inconceivable that the nervous system could affect cellular oxidations directly. To these assumptions Wielowiejski added the very dubious argument that the reason why an isolated organ can glow for a long time is that excess photogenic material is formed during the act of extirpation. Bongardt advanced what seems to be the diametrically opposed argument that photogenic material must be stored in excess, because fireflies are not able to extinguish their light suddenly. It appears, however, that both these workers had vaguely in mind the concept that luminescence continues as long as stored substrate remains, but that they differed in their ideas of how long luminescence continued and how quickly it could be terminated. Their accounts are further confused by failure to distinguish the luminescence of dead fireflies from that under normal control, and by the fact that their evidence against oxygen as a controlling factor is very equivocal.

Gerretsen's views on reactant limitation suffer from the same sort of ambiguous duality, since he assumed control by the end-cell (oxygen limitation) but, at the same time, postulated that stimulation produces a substance which emits light when it comes in contact with oxygen (control by substrate synthesis). However, Gerretsen did elaborate on the relation between synthesis and luminescence, in postulating, by analogy with Harvey's early chemical work, that the "oxidized light material" was reduced during the dark period between flashes and thus became ready for luminescence again. This idea had no experimental support until very recently, when Alexander, on the basis of his work on the effects of high oxygen tensions, postulated that luciferin "is steadily accumulating in a potentially reactive state within the photogenic cells between flashes", and that under circumstances when oxygen ceases to be limiting (*i.e.*, is present in excess), "the rate of luciferin release . . . becomes the critical factor for luminescence." Alexander, however, considered luciferin release of importance only in the hyperoxic glow and not in the control of flashing. He was unable to define precisely what should be understood by "luciferin accumulation" and "luciferin release." If either process involved synthesis, it would be hard to reconcile with continuing luminescence after death. On the other hand, if "release" were to mean a sort of reduction*, depolymerization, or hydrolysis of stored reserve, Alexander's results might be useful in interpreting long-sustained luminescences. It would, of course, have to be assumed that the "release" could continue in fireflies dried and remoistened, and that sufficient luciferin could "accumulate" prior to death to last the required time.

Space does not permit a detailed analysis of the data from which Alexander drew his conclusion. There were, however, a number of lines of evidence which fitted together consistently. On the other hand, to deal with the control of ordinary flashes and glows, Alexander adopted intact the end-cell theory of Dahlgren, which will be discussed later. It is, therefore, a possible weakness in Alexander's work that two entirely different mechanisms have to be invoked to account for all the luminescent phenomena observable in the firefly. All of Alexander's data, moreover, can be explained consistently by the modified mechanical end-cell theory which I shall present in the last section of this paper. Again, space does not permit a point-by-point comparison of the two schemes at this time. In any case, no final decision could be reached as to their respective merits, since both fit the available evidence satisfactorily. However, I consider it an advantage of the mechanical end-cell theory that it explains both normal flashing and glowing, and also the various abnormal types of luminescence.

In summary, then, there are two views in regard to the possibility that luciferin (or luciferase) could be the limiting factor in the control of luminescence. By far the most popular view is that a large excess of stored material is always present in the photogenic organ. This is supported by the common observation of long sustained luminescences, particularly in dead animals. The other view is that the substrate for luminescence is being synthesized constantly, and that there is never enough of it on hand for more than one flash. As we have seen, the evidence for this view, in the form stated, is not strong. If, on the other hand, the limitation is thought of in terms of luciferin *release,* as suggested by Alexander, rather than luciferin *synthesis,* there is a possibility of reconciling the two views, in the sense that stored reserves could exist. It should be kept in mind, however, that it would still be necessary to exclude the possibility of oxygen limitation, and to show that luciferin accumulation (synthesis) and release could explain all normal luminescent phenomena.

---

* See remarks at the end of the section on the reflector layer (page 408).

Oxygen as a Possible Limiting Factor in the Control of Luminescence.* Although depriving fireflies of oxygen in various ways has been perhaps the most popular single experiment on bioluminescence, the instances are rare when this has been performed or described with sufficient precision to be of much value. To the familiar difficulties of failure to distinguish abnormal from normal luminescence, and effects on control systems from effects on luminescence *per se,* there has been added the failure to use strictly oxygen-free gases and leak-proof exposure chambers. In addition, sufficient attention often has not been paid to the facts that many common fireflies rarely flash or glow spontaneously in captivity, and that practically all fireflies are normally non-luminous by day (many observers) and in bright light (Buck, 1937a, and others). Hence, extinction of light in a given gas, or failure to luminesce, is not in itself necessarily proof that luminescence cannot occur in that gas.

In spite of these experimental deficiencies, there is no possible doubt that oxygen has a profound role in photogeny in fireflies. Luminescence has been reported to be reversibly extinguished (eventually), or not to develop, in pure N_2, H_2, or CO_2 (Forster, 1783; Spallanzani; Grotthuss, 1807; Macaire; Matteucci, 1843; Joseph; Owsjannikow, 1864; Bellesme, Dubois, Watasé, Bongardt, Townsend, Shafer, 1911; Kastle and McDermott, Creighton, Snell, Emerson, 1935; Hasama, 1942a; Alexander; Buck, 1946b). Snell's results are particularly valuable, since he was able to show that luminescence is possible below 4 mm. of O_2. This astonishingly low oxygen requirement suffices to explain reports of failure to get extinction of light with N_2, H_2 or CO_2 (Davy, 1810; Macartney, Bongardt, Kastle and McDermott). It also accounts for continued luminescence under oil (Carradori) or boiled water (Dubois), although utilization of air trapped in the tracheae may also have been involved (see p. 444). It also suggests that a control mechanism which operated by regulating oxygen would have to be of extraordinary efficiency. Additional indication of the necessity for oxygen is furnished by the observation that luminescence disappears in an evacuated space (Macaire, Owsjannikow, Dubois, Bongardt, Snell, Alexander). Here, the failures of Carradori, and of Kastle and McDermott, can be ascribed to insufficient pressure reduction, as can Dubois' claim that a dried organ glows in a vacuum if moistened.*

---

* In this paper, "anaerobic" and "anoxic" will be used in their strict meanings of absence of oxygen, while "hypoxic" will be used for tensions (partial pressures) of oxygen (pO_2) which are abnormally low.

Macaire made the interesting observation that fireflies made non-luminous in a vacuum cannot be induced to luminesce, by subjection to either heat or electricity. Similar observations were made later by Bellesme and by Knoche (1910) in regard to electrical stimulation of fireflies rendered non-luminous in irrespirable gases. These experiments have been widely interpreted as proving, first, that there can be no direct nervous stimulation of the photogenic cell; and, second, that the normal control of luminescence is by way of oxygen. Since these are very important conclusions, it is necessary to emphasize that the abovementioned experiments are entirely vitiated by the probability of an anoxic effect upon the nerves themselves. In this connection, it is interesting to note that the voluntary control mechanism is usually inactivated by oxygen lack before luminescence is abolished (see pp. 444 *et seq.*). Moreover, in regard to the relevance of these experiments to direct nerve action, there is no reason to expect that luminescence could occur in the absence of oxygen.

Arnold (1881) claimed that electrical stimulation induced *L. noctiluca* to light in "absolutely oxygen-free hydrogen" (as judged by the disappearance of the glow of stick phosphorus in the same chamber). However, since (a) he stimulated only within five minutes of the darkening of the phosphorus, (b) the glow persisted for a time after the current was broken, and (c) he could get only a single response, the glowing was probably made possible by air still remaining in the interior of the tracheal system, as suggested by Edwards (1863).

Experiments with increased oxygen have also been tried, but with conflicting results. Augmented luminescence has been observed by Forster, Spallanzani, Matteucci, Owsjannikow, Bellesme, Severn (1881), Kuhnt, and Emerson and Emerson (1941), no change by Hermbstadt (1808), Davy, Kölliker, Macartney, and Dubois; and an actual decrease in emission by Macaire, Bongardt, and Hasama. Bongardt and Hasama claimed this last result to be a purely secondary one caused by the inhibitory effect of the actual gas current, because a stream of ordinary air likewise inhibited luminescence. However, their contention that this is also the explanation of the effects of N_2, H_2, and CO_2, is logically unsound. Moreover, Knoche, and Höllrigl, in repeating Bongardt's work, showed that when precautions were taken to remove O_2 from CO_2 and H_2 completely, no luminescence could be elicited in *Lampyris*. Knoche also found a stimulating rather than depressing effect of gas flow. I have reported a similar effect of gentle air-currents on the flashing of fireflies in flight in the laboratory (Buck, 1937a). Severn reported that air currents had no effect on luminescence. However, since inhibitory effects of strong air currents have been observed in the field by Geipel, and by myself, and in the laboratory by Dubois, a reflex mechanical inhibition of lighting, in accord with Bongardt's observation, is a possible source of error to be kept in mind in all work on gases, since their effects are almost always tested in streams. The effect of elevated oxygen tension is also seen in the increased luminescence obtained with high air pressure (Dubois, Lund, Alexander). According to Alexander, this corresponds with the effect obtained with nitrogen-oxygen mixtures at atmospheric pressure but with equivalent partial pressures of oxygen. Heinemann (1886) may have been dealing with the same phenomenon when he found that blowing air through a thoracic spiracle of *Pyrophorus* increased the light, although mechanical stimulation is a more likely explanation. None of the above effects bears on whether or not the oxygen effect is direct or indirect, since the experimental conditions may well have been stimulating to nerves as well as to photogenic tissue.

---

* There are, as a matter of fact, two possible objections to results obtained with vacuum. First, if evaporation of water were fast enough, the cooling effect on the firefly might in itself inhibit luminescence. Second, if sufficient water were lost, lighting would be limited by water, rather than by oxygen. Probably neither objection is serious in short-term experiments with moderate rates of evacuation.

Kastle and McDermott, Snell, Maloeuf (1938), and Alexander, in work on the effect of increased oxygen on *Photinus pyralis* and *Photuris pennsylvanica,* showed that these fireflies may develop a steady glow, the mechanism of which will be discussed later ("hyperoxic glow").

Another line of inquiry which bears on oxygen limitation is one undertaken by Snell, at the suggestion of Harvey. Snell attempted to ascertain whether the nervous control of luminescence is direct or indirect, by recording flashes of *Photuris pennsylvanica* in different oxygen tensions. The analysis was based on two assumptions: first, that the actual luminescent reaction is independent of pO_2 above about 2 mm., as Shoup (1929) had found to be true in luminous bacteria; and second, that the control mechanism opens to the same extent and for the same period each time it is stimulated. It was then argued that, if the amount of light in the flashes were independent of oxygen tension over a wide range, it would indicate that luminescence was controlled by "removal of a phase boundary between reactants" in the photogenic cytoplasm, due to direct nervous stimulation. Dependence on oxygen, on the other hand, would indicate that an end-cell valve was operating, because a given response would admit less oxygen at low tensions and thus support less total light-emission. Snell did not actually compare total emissions, but only the maximum intensities of the flashes at various oxygen tensions. He found that there is apparently a progressive, though moderate, decrease in peak intensity between 150 and about 20 mm. oxygen. From this, he concluded that luminescence is controlled by regulation of oxygen by the end-cells.

It seems to me that Snell's conclusion is not justified. In the first place, there is no reliable evidence that the end-cell is the actual control mechanism. Secondly, since the constancy of operation of the control mechanism is a pure assumption, it could as well be argued that the results were caused by a slight progressive change in the control mechanism, such as might be induced in the nervous system by the changing pO_2. As a matter of fact, since Snell observed only about a 40 per cent decrease in average maximum intensity over a 7-fold change in pO_2, there must have been very considerable compensatory changes in the response of the "end-cells", if luminescence *per se* is to be considered as having been influenced by oxygen in the range from 150 to 20 mm. In general, then, the only conclusion justified from this work is that the intensity of luminescence varies with oxygen tension in the range explored. Direct nerve action is by no means excluded.

In summary, most of the careful work agrees in making oxygen essential for continued luminescence. Under various experimental conditions, oxygen may be the limiting factor in the control of photogeny, but there is no direct evidence that oxygen is ever limiting in normal flashing. Nevertheless, there is sufficient circumstantial evidence of normal control by oxygen to justify the additional inquiries described on the following pages. Thus far, the evidence has not permitted a decision as to whether the presumed control comes about by an actual limitation of oxygen as a reactant, or secondarily by the effect of oxygen limitation on some other mechanism such as the nervous system.

The Spiracular Factor in Luminescence Control. Since the spiracles guard the primary site of entry of oxygen into the tracheal system, and since, in many insects, they are known to be efficient in barring access of gases or vapors (Wigglesworth, 1939; Hoskins), their possible influence on luminescence in the firefly invites investigation. Maloeuf and Alexander mentioned the spiracles briefly in connection with their theories, but apparently they made no observations. On the other hand, Wielowiejski, Lund, and Gerretsen all argued that no control located in the large tracheae could account for the rapidity with which a flash starts and stops, because of the time required to exhaust the oxygen in the tracheal system distal to the valve. In agreement with this view, I found no correlation between the state of the spiracles and the time of occurrence or characteristics of normal or electrically induced flashes in *Photinus pyralis* (Buck, 1946b).

In the same note, I reported investigations on the influence of the spiracles on glows produced by anoxia, ether, cyanide, distilled water, and adrenalin. Injected water caused immediate spiracular closure in both normal fireflies and those glowing from ether or cyanide, and subsequent cessation of glow in the latter. With all the other treatments, the spiracles opened before glowing began, and in anoxia they closed after the glow ceased. They usually also closed at about the time the glow ceased in ether or cyanide vapor, or with injected adrenalin, but not infrequently the glow continued for a few minutes after the spiracles had closed. This was construed as showing that an "internal control" was still "open." Calculations based on measurements of the tracheae immediately associated with the light organ showed that a sufficient volume of air is trapped in and close to the light organ to support luminescence for several minutes, if the oxygen requirements of the light organ are comparable to those of ordinary tissues.

There is, apparently, a further stage of spiracular opening following the secondary closure, because dead or moribund fireflies usually have their spiracles open, whether glowing or not.

From this work, it was concluded that, although the spiracles and "internal control" often operate simultaneously, the spiracles do not ordinarily exercise any direct control over induced glows in flashing types of fireflies. However, the calculations indicate that they can terminate glowing eventually, and I think it not unlikely that they are important in luminescence control in normally glowing types of fireflies such as *Diphotus.* I attempted to investigate this question in the larva of *Photuris pennsylvanica,* but found the spiracles too small.

As argued in regard to breathing movements, if the spiracles were the only means of oxygen control for the organism, luminescence would be expected to occur by day, since the spiracles would presumably have to open to meet respiratory needs. This would not agree with the usual reports that fireflies are non-luminous by day. However, since fireflies are also inactive by day, it is possible that their general metabolism falls so low that the spiracles would need to be opened only very infrequently. Under such circumstances, even if luminescence occurred when the spiracles opened, it might escape observation.

The Internal Control of Glowing. A large number of agents and treatments have been discovered which cause glowing in Type 6 fireflies (Verworn, Kastle and McDermott, Gerretsen, Creighton, Snell, Emerson, Emerson and Emerson, Alexander; Seifter, 1945; etc.). Of these, hypoxia and anesthetic vapors have proved to be particularly useful in studying the operation of the control mechanism, and from their use a number of conclusions have been drawn concerning the tracheal end-cell. Since none of this work actually bears on the anatomical identity of the regulatory mechanism, I shall use the noncommital term, *internal control.*

Verworn, and also Gerretsen, studied the effect of strong chloroform vapor on *Luciola italica* and *Luciola vittata,* respectively, forms which apparently give frequent spontaneous flashes under laboratory conditions, in contrast to the rare flashing of our American *Photinus pyralis.* They observed three stages of chloroform effect: (1) reversible inhibition of flashing, with immobilization, and sometimes accompanied by a dull glow; (2) irreversible bright glowing; (3) irreversible extinction. Stage 1 was ascribed by Verworn to nerve narcosis, and by Gerretsen to transient contraction of the end-cells. Stage 2 was assigned by Verworn to direct chemical stimulation of the nerves and photogenic tissue, and by Gerretsen to relaxation of the end-cells. Stage 3 was attributed by Verworn to exhaustion of the stored photogenic material, due to the luminescence in the second stage, and by Gerretsen to permanent damage to enzymes. Verworn's opinion in regard to Stage 3 is weakened by Gerretsen's observation that the dead fireflies in Stage 2 sometimes glowed for hours if removed from the chloroform vapor. The idea of some sort of irreversible chemical damage to the luminescent system is supported also by my observations on *Photuris pennsylvanica.* In this species, I found that light organs from specimens rendered non-luminous by an hour's exposure to saturated ether vapor, would not glow even if minced in fresh air, a treatment which ordinarily suffices to cause glowing even in moribund animals. The above idea is also supported indirectly by Emerson's work with known concentrations of ether, in which he found that irreversible damage to the luminescent system results when lethal concentrations are reached.

Either because they used saturated vapor or because chloroform is rather toxic, Verworn and Gerretsen missed some important intermediate stages of anesthesia which can be demonstrated in gradual hypoxia and with gradually increasing concentrations of ether. Thus, Emerson, and also Buck (1946b) have shown that bright early ether glows are readily reversible. The same is true for glows produced by low oxygen (Bongardt, Kastle and McDermott, Snell, Alexander; Buck, 1946b).

Snell studied the effect of low oxygen tension carefully, and showed not only that partial pressures below 4 mm. induce a steady dull hypoxic ("anoxic") glow* but also that, if the oxygen tension is suddenly raised from below 4 mm. to 30 mm. or above, the glow brightens suddenly, and usually markedly, and then quickly subsides to zero ("pseudoflash"). Snell interpreted the hypoxic glow as due to entry of oxygen through inactivated end-cells, and the pseudoflash as an extra uncontrolled luminescence which is terminated by recovery and closure of the control mechanism (end-cell valve).†

A further refinement in "internal control" action was disclosed by my finding (Buck, 1946b) that with very gradually decreasing pO_2 a stage is reached where an hypoxic glow has not yet developed, but in which a pseudoflash can be induced. To explain this phenomenon, I postulated that the "aperture" of the internal control had enlarged enough to permit luminescence at high pO_2 (pseudoflash), but still not enough to permit it at low. Since luminescence (hypoxic glow) does develop at oxygen tensions still lower than in the above stage, I concluded that the aperture of the internal control enlarges faster than the pO_2 decreases. In other words, the internal control overcompensates in its response to decreasing oxygen tension. These facts enable us to make the further deductions that luminescence requires the entry of a certain minimum number of oxygen molecules per unit time, and that this can be achieved either by the entry of air higher in oxygen through a smaller aperture, or by the entry of air lower in oxygen through a larger aperture. This implies, furthermore, that the response of the internal control is a graded, rather than all-or-none phenomenon.

---

* The hypoxic glow is ordinarily produced by using pure N_2, H_2 or CO_2, but Bongardt, Snell, and Alexander showed that it occurs also in atmospheric air if the total pressure is lowered enough to decrease pO_2 below 4 mm.

† The pseudoflash has been interpreted by Alexander in terms of the Wigglesworth-Maloeuf osmotic theory, which will be discussed later. It might also represent the oxidation of luciferin that had accumulated during the hypoxic period, as has been shown to occur in the "flash" of luminous bacteria. However, in a number of experiments with *Photinus pyralis,* I found no correlation between the durations of pseudoflashes and the durations of the preceding anoxic periods.

Further information on the operation of the "internal control" is contained in a number of observations bearing on whether or not flashing and glowing are controlled by the same mechanism. It has been reported generally that normal flashing does not occur under anesthesia, except occasionally at the start of exposure, before glowing has developed to full intensity. Unfortunately, the evidence is not clear-cut on whether or not flashing can be elicited artificially during glowing. Although his account is very ambiguous, Gerretsen apparently believed that luminescence could be affected by electrical shock during his Stage 2 (glowing) of chloroform narcosis. On the other hand, Snell (1932) and Alexander found that normal flashing is inhibited during the hypoxic glow. Furthermore, Alexander was unable to induce flashing during hypoxic glowing, by electrical treatment. These results have been regarded by Alexander, and others, as showing that the flash control cannot operate during glowing, or in other words, that flashing and glowing are independent phenomena. However, none of this work circumvents the objection that that hypoxia and anesthesia (like anoxia) undoubtedly abolish the normal functioning of the flash control mechanism. Furthermore, the failure to obtain flashing with hypoxia might have been due simply to the very low pO_2, rather than to a necessary difference in the control mechanisms for flashing and glowing. In other words, luminescence may already have been maximal for the existing pO_2.

Snell's work with low oxygen is particularly germane with regard to the distinction between glow and flash. He found that, below about 20 mm. oxygen tension, the duration of the flash lengthened progressively and the peak intensity decreased until, at 4 mm., the steady hypoxic glow resulted. The region between 20 and 4 mm. oxygen tension, therefore, is a transition territory in which it is hardly possible to distinguish flash from glow.

An additional point of importance concerns the normal "neuromuscular" state of the internal control mechanism. The fact that fireflies moribund from inanition (or those which fail to recover from moderate anesthesia) often exhibit a steady, lasting glow, indicates that in its relaxed position the "valve" is open, and that energy expenditure is required to keep it closed. The same conclusion is indicated by the work of Gerretsen and of Emerson on anesthetics, and by the work on the effects of oxygen, if the assumption is made that the internal control relaxes (opens) in low oxygen and contracts (shuts) in high (Snell, Alexander; Buck, 1946b). As a matter of fact, adrenalin, cyanide, and all the other glow-inducing agents can be imagined to have an analogous effect. The open position in hypoxia, anoxia, anesthesia, and death is compatible with an assumption that energy expenditure is required to keep the "valve" in the closed (tonic), or non-luminescent, condition. This point will be discussed further on page 450.

A further interesting point was discovered by Alexander in his extension of Snell's work on the effects of high oxygen tensions. Although there is not space here to give the evidence in detail, Alexander was led to conclude that the glow produced by high oxygen is due to oxygen leaking directly through the cytoplasm, thus by-passing the internal control.

In experiments on glows produced by ether, carbon dioxide, and adrenalin, I found the responses of *Photuris pennsylvanica* larvae similar to those of the adult *Photinus pyralis.* Since the larva lacks end-cells, it can be concluded that end-cells are not necessarily involved in the control of glow in adult fireflies. However, they might still be essential to flash control, since none of the work discussed in this section really touches that point.

Theoretical Mechanisms of Luminescence Control by Oxygen Limitation:

INTRODUCTION. In this section, the mechanisms of oxygen control which have been proposed in sufficient detail will be analyzed in the light of the anatomical and physiological evidence discussed, and according to certain theoretical considerations. A review of the anatomical path which oxygen must follow in its journey to the photogenic cells, suggests the following sites at which the passage of oxygen might be interrupted: the spiracle, the end-cell, the tracheole, the differentiated zone, the membrane of the photogenic cell, and the photogenic cytoplasm. The spiracle has already been considered and found not to be concerned in the control of flashing. What little information exists on the differentiated zone of the photogenic cytoplasm, and on the possibility of phase-boundary changes in the cytoplasm, has been presented and need not be reconsidered. The membrane of the photogenic cell is theoretically the most logical site of all for the operation of an efficient control, but since Harvey (1922) has shown that a number of cells are very permeable to oxygen, there is no reason to expect that control by changes in permeability to oxygen is likely.* We are reduced, therefore, to the tracheoles and the end-cells as possible agents in oxygen control.

---

* There is every reason to believe that this conclusion is valid generally, but Harvey's work actually concerned permeability under abnormal conditions (recovery from anoxia).

The discussion will be mainly in terms of the Type 6 organ, since it has been the subject of nearly all the modern physiological work. However, this should not justify any generalizations, either pro or con, concerning any particular theory. Thus, it seems certain from the time relations, if from nothing else, that a firefly with an intermittently glowing type of organ will require a very much less precise and intricate control mechanism than one of Type 6.

COMPETITION FOR OXYGEN, IN RELATION TO CONTROL. One question which may have an important bearing on control is whether or not there is competition for oxygen between respiration and luminescence. This possibility exists because both systems presumably draw their oxygen from a common intracellular "pool".*

There would be potential competition if, during luminescence, the oxygen supply fell below that necessary to support both luminescence and basal respiration. There are few data which bear, even indirectly, on what occurs in these abnormal circumstances. Harvey (1922) found that, in *Chaetopterus* sealed in water, luminescence disappeared long before spontaneous muscular activity. However this may not indicate that luminescence is more sensitive to oxygen lack than is respiration, but only that the oxygen uptake necessary to complete the enzymatic reactions in muscle can be long delayed, whereas those in luminescence cannot. Moreover, it is possible that the effect was due to accumulation of carbon dioxide, rather than to deprivation of oxygen. Shoup found that respiration† of luminous bacteria began to fall off when pO_2 had been reduced below 23 mm., whereas luminous intensity was unaffected until pO_2 reached 2 mm., at which point respiration was only 50 per cent of normal. If, as seems reasonable, the intensity of luminescence is proportional to the oxygen consumption of the luminescent reaction, these results suggest that luminescence may be more successful in obtaining oxygen from a deficient supply than is respiration. It is, however, not certain that this conclusion is applicable to fireflies.

If, as Shoup's results suggest, luminescence has a competitive advantage over respiration under conditions of restricted oxygen, it appears unlikely that luminescence could be suppressed by respiration under normal circumstances, where sufficient oxygen for respiration is always present. However, since Creighton has proposed such a mechanism, it will be worthwhile to analyze it in some detail, particularly as the analysis points to a possible experimental test.

---

* Recent work on luminous bacteria indicates that, although respiration and luminescence show considerably different responses to certain stimulating and inhibiting agents (*e.g.*, potassium cyanide and urethane), they utilize parallel enzymatic pathways. With increased oxygen, for example, proportionately more oxygen goes through the luminescent system. (See Harvey, 1940, 1941, and Johnson et al.)

† Shoup used "respiration" as synonymous with total oxygen uptake, but in the present paper it is used in the restricted sense of oxygen uptake other than that required for luminescence. In practice, the two measures will probably be nearly identical, since Shoup's work, and that of Snell on the hypoxic glow of the firefly, indicate that the oxygen requirement of luminescence is very small in comparison with that of respiration.

In view of Shoup's work, respiration could deprive luminescence of oxygen only under one of the following conditions: (a) the oxygen requirements of respiration are actually lower than those of luminescence, and the photogenic tissue is maintained nearly anaerobic, at least during the dark periods; (b) respiratory activity is localized in a cortex surrounding the actual photogenic cytoplasm, so that oxygen diffusing in it can be "filtered" out before reaching the other reactions of luminescence; (c) the intracellular oxygen transport system channels oxygen preferentially into respiration (at normal pO_2). The first possibility is rendered very unlikely by the known low oxygen requirement of luminescence. The second possibility certainly appears unreasonable on general grounds. However, in its light, Dahlgren's speculations on the "oxygen-insulating" properties of the differentiated layer of the photogenic cytoplasm (see p. 413) assume a new and provocative significance, and perhaps deserve further consideration. The third condition is one which, by analogy with known systems, is possible, but which would not be expected to permit absolute extinction of luminescence,* since such "channelings" are usually questions of relative rates of utilization, rather than absolute exclusion.

Since the idea of competition seems not impossible, an enumeration of its nature and consequences is in order. According to Creighton's hypothesis, basal respiration normally uses all the oxygen available and, hence, prevents visible luminescence (dark period). If, however, additional oxygen is made available (*e.g.*, by end-cell contraction), photogeny may begin (luminescent period). There is, however, an alternative method by which competition could operate. According to this second hypothesis, the oxygen supply remains relatively constant, but is usually all appropriated by respiration, which proceeds at a rate higher than basal (dark period). If, however, the respiratory rate falls, oxygen is available for the luminescent reaction. Both hypotheses, it will be seen, are compatible with the fact that respiration can be maintained without permitting photogeny at "inappropriate" times. This is particularly pertinent in view of the very small pO_2 required for luminescence. Both hypotheses also are in harmony with the expectation that respiration is diminished or absent in moribund or dead fireflies, where luminescence is often continuous. Assuming that prolonged anesthesia depresses respiration, both hypotheses also explain glowing in narcosis. Neither hypothesis accounts for either hyperoxic or hypoxic glowing, without further assumptions which are too detailed to consider here.

Theoretically, a decision as to which mechanism (if either) is operating could be made by comparing the oxygen uptakes of a tissue in the dark and in the luminescent states. According to the first hypothesis, oxygen consumption should be higher during luminescence than during the dark period; according to the second, probably lower.[*] Unfortunately, as already discussed, no direct data are available.[†] However, it will be instructive to pursue some of the consequences of the two hypotheses a little further.

---

* In view of the extreme sensitivity of the human eye, it is likely that luminescence really is zero during the apparently "dark" period, rather than continuing at a low rate.

The first hypothesis, according to which respiration is lowest during the dark period, fits better the traditional idea that control is achieved by limitation of the external oxygen supply. It agrees, for example, with the expectation that luminescence involves increased oxygen and that nervous stimulation induces luminescence by inducing an increase in that supply. Furthermore, a linear relation between luminescence and oxygen tension, at least over part of the range, would be expected on the basis of the first hypothesis, if the oxygen uptake of luminescence is linear with oxygen tension. No relevant work has been done on glowing forms, which would be the most favorable for the purpose; however, as we have seen, Snell's data indicated that peak flash intensity decreases as the partial pressure of oxygen decreases from atmospheric. On the other hand, in several organisms (notably bacteria), luminescence has been found to be independent of oxygen over very wide ranges.

The second hypothesis, in common with other intracellular controls, has the advantage over Creighton's mechanism of potentially greater speed, and the apparent "disadvantage" of not accounting for the end-cell. In addition, it requires us to think of nervous "stimulation" of luminescence in terms of inhibition of respiration, and of the firefly as expending more energy in keeping itself dark than in luminescing.

On the whole, although a competitive mechanism is not excluded from being concerned in firefly luminescence, the probability does not seem large enough to justify pursuing the question further. Likewise, in the absence of experimental work, there seems to be little point in attempting to choose between Creighton's hypothesis and its alternative.

The Theory of Osmotic Control of Luminescence. In 1930, Wigglesworth proposed an ingenious theory, according to which the changing needs of a cell for oxygen could be met by varying the amount of fluid in the distal end of the tracheole. This theory was not originally designed for the special anatomy of the firefly, with its end-cells and anastomosing tracheoles and, moreover, it assumed intracellular penetration of the tracheoles. Several years later, however, it was applied specifically to the firefly by Maloeuf.

---

[*] It is generally assumed that the oxygen consumption of the luminescent reaction occurs at the same time the light is emitted. This appears to be supported by the fact that oxygen is essential for continued luminescence. However, over periods as short as those of normal flashes, it might well be that light-emission and oxygen uptake would not be simultaneous. An analogy is offered by muscle, in which the oxygen is used in the recovery or preparatory period, whereas the actual reactions involved in contraction occur anaerobically. Therefore, a comparison of oxygen uptakes during light and dark states of a tissue might not give a true measure of the extra oxygen consumption involved in the luminescent reaction.

[†] Some idea of the oxygen requirements of the same tissue in the luminous and non-luminous states might be obtained by comparing quiescent normal fireflies with brightly glowing "dead" specimens, or, perhaps (to eliminate the possibility of oxygen uptake by surviving tissues or by bacteria), with fireflies which had been dried and then remoistened. This would not, however, be ideal, because basal respiration would be lacking in the dead group. It may be of interest, also, to mention here my preliminary experiments on the oxygen uptake of specimens of *Photinus pyralis* poisoned with DDT. As with other insects, oxygen uptake was increased several-fold. Luminescence was also, at first, increased over the near inactivity usual in the laboratory, but consisted of irregular flashing and intermittent glowing, suggestive of nerve irritation. It was only late in the experiment, when muscular activity and oxygen uptake had both declined, that luminescence became continuous.

Wigglesworth proposed, first, that the tracheolar wall is semi-permeable. Second, he postulated that the level to which air extends in the tracheoles is determined by a balance between the osmotic pressure of the cytoplasm (which tends to pull water out of the tracheole into the cell) and the capillary attraction of the liquid in the tube (which tends to draw water out of the cell). He supported this idea by showing that hypertonic solutions injected into mosquito larvae caused movements of the air in the tracheoles. He further assumed that, if the osmotic pressure into the cell were increased during cell activity, as by the breaking-down of large substrate molecules into smaller metabolic products such as lactic acid, water would be withdrawn from the tracheole into the cell. This would allow the air to extend further in toward the cell and meet the increased need for oxygen imposed by the increased activity. Conversely, during quiescence, materials would be resynthesized in the cytoplasm, causing its osmotic pressure to fall, and water would pass from the cell into the tracheole and force the oxygen supply farther away. In agreement with expectation, Wigglesworth was able to show reversible inward movement of air, and acid formation, during anoxia. Maloeuf's contribution consisted in defining the "activity" which increases the osmotic pressure, as the luminescence of the photogenic cell, and in showing that injection of hypertonic and hypotonic solutions into the photogenic organ caused, respectively, stimulation and suppression of luminescence. Dubois, however, had found that water injection induced glowing.

The osmotic control theory has run into several difficulties. First, Wigglesworth (1938 a and b) found, in developing mosquito larvae, a number of phenomena which were not explicable by his theory in its original form. Shortly thereafter, Bult (1939) introduced a considerably altered version of the theory and, by processes too complex to elaborate here, came to the conclusion that an increase in osmotic pressure in the cell would actually bring about movement of air *away* from the cell. The "osmotic" effects he therefore attributed to swelling of cell proteins. The Bult theory could probably be adapted to the problem of photogenic control, although it is by no means clear that it is a substantial improvement over the original Wigglesworth theory. Both theories, however, suffer from the serious drawback that the observed speed of the fluid movements produced is very slow. Over distances comparable to those met in the tracheoles of the light organ, the time required is of the order of minutes, where seconds or fractions of seconds are required.

Alexander repeated Maloeuf's experiments in detail and confirmed the findings in regard to hypertonic solutions, but not concerning hypotonic solutions. In some instances, moreover, he found that the direction of spread of glow in the organ did not correspond with that expected by theory. He concluded that tracheolar fluid movements are not concerned in normal flashing, although they probably are in the glow caused by injected hypertonic solutions, and in the fact that the pseudoflash has a shorter latent period than has hyperoxic glow. Since we have already seen that many substances beside hypertonic solutions induce glowing, this effect cannot safely be ascribed to osmotic changes. Likewise, as will be shown shortly, the length of the latent period between stimulus and light-production can be explained adequately by a valve mechanism which opens in low oxygen tensions and closes in high. Another explanation of the effect of fluid in the tracheoles will also be presented.

The Mechanical Valve Theory of End-Cell Control of Oxygen. By far the majority of theories concerning the end-cell give it the role of a regulator of oxygen rather than a source or adjunct of luminescence. Most of these theories are in terms of the generalities which have already been discussed, as, *e.g.*, the strategic location of the end-cell, its supposed innervation, and especially its presence in flashing fireflies and absence in glowing ones.* This last point seems to me to be a particularly strong one, though, admittedly, it provides no information on how the end-cell works. The objection which is often raised that end-cells are found in non-luminous tissues, even in the firefly, does not show that they cannot operate in regulating oxygen. Schultze has pointed out that it is not the exclusive possession of end-cells, but their strikingly high concentration, which distinguishes the photogenic layer from other tissues.

Dahlgren and most other writers on end-cell physiology have assumed that the tracheoles are at least partly air-filled. As we have seen, the evidence is by no means unequivocal in this respect, but unless it is assumed that an actual current of liquid is circulated through the tracheoles, we shall have to postulate the presence of gas distal to the end-cell, as well as proximal to it, if the end-cell is to operate as a valve. This is because the speed of diffusion of oxygen through tissue is not greatly different from that through plain water (Krogh), so that, if the tracheoles were water-filled, the end-cell would offer relatively little hindrance to aqueous diffusion.

Though a number of investigators have thought of the end-cell in frankly physical terms, Dahlgren is almost alone in describing structures which might make possible the valvular action. As already mentioned, he found, in osmic acid preparations, that the tracheal twig narrows within the end-cell, loses its spiral thickenings, and shows an annular "dark-staining sheath", "specialized body", "cylindrical organ" or

---

* Maloeuf has made much of his claim that Bongardt found end-cells in the larva of *Phosphaenus hemipterus*, an animal which, according to Maloeuf, "displays no such brisk flashing rhythm." I would not regard this as conclusive in any case, in view of the contradictions in Bongardt's paper but, actually, Bongardt (p. 24) says that end-cells are absent.

"rounded mass" ("S" in FIGURE 11).* This staining reaction, continues Dahlgren, "may point to a contractile layer of cytoplasm surrounding it, with the possibility of a valve, in addition, in its lumen. Also the larger body of cytoplasm surrounding it sometimes shows a radial structure that may point to a general contractile power."†

Dahlgren explains the functioning of the end-cell as follows: "It is believed by the writer that these radiating rods and the rounded mass represent muscular structures which are of two possible uses: to prevent the passage of air into the tracheole by compressing them; this would appear to be the function of the rounded mass; second, to enlarge the terminal twig and end organ and thus fill it with a new supply of air; this would seem to be a possible function of the radiating rods. It can be seen that the rapid alternation of the two actions would result in the forcing of a sudden jet of air bearing free oxygen into the light cell mass, and would account for the power of flashing exhibited by the organ." Dahlgren suggests, in addition, that two sorts of nerves will be required, one for the rounded body and one for the fibrillar cytoplasm. His hypothesis was used uncritically by Creighton, Snell, and Alexander in their physiological work. Alexander, in fact, interprets most of his results on the basis of end-cells able to cause a "sudden burst of oxygen directly into the photogenic cells".

It seems to me that Dahlgren's mechanism, in the form stated, is equivalent to pinching a tapering tube open at both ends, and expecting the jet of air to move toward the smaller end (that is, into the tracheoles). Matters would be even worse if the small end of the tube were partly or entirely closed, as it might be if the tracheole contained some water or terminated intracellularly. In order to enable Dahlgren's end-cell to actually produce a "jet" of air, several additional mechanical properties must be postulated for it, and a number of new assumptions introduced. Since these requirements apply to any mechanical end-cell valve theory, it is worth while to enumerate them, as follows.

(a) The "rounded mass" must lie proximal to the region of the end-cell lumen which is enlarged by contraction of the fibrillar protoplasm. This is necessary in order that the contraction of the rounded body can shut one end of the tube through the end-cell, so as to allow pressure to build up distally.‡

(b) The passage through the end-cell must have an elastic wall, in order to provide a means of applying pressure to the air which is drawn into the end-cell when the passage is enlarged by contraction of the fibrillar cytoplasm.

---

* Dahlgren seems to be somewhat vague about the exact location of this body, since in his Figure 11 (p. 611) he shows it as indicated in the lower of the two end-cells pictured in my FIGURE 11, whereas in another version (his Figure 20, p. 345), diagrammed as the upper of the two end-cells in my FIGURE 11, it does not seem to appear.

† I, too, have seen a striated appearance in the end-cell cytoplasm in some preparations (Buck, 1940), but the structures are not very sharp. They, and indeed the whole end-cell, are very small.

‡ Dr. H. Specht has pointed out to me that the internal surface of the end-cell lumen must be kept dry, since otherwise a meniscus would form at the point of constriction, and resist, with relatively enormous force, the reopening of the tubule.

(c) The periphery of the end-cell must be rigid. This will be necessary in order to give the "fibers" something to anchor to as they pull on the wall of the lumen of the end-cell. As already mentioned, the periphery of the end-cell is notably lacking in any suggestion of such a compact structure.

(d) By analogy with vertebrate sphincters, the "rounded body" ought to be composed of two layers of contractile elements at right angles to one another, one for enlarging its opening and one for constriction.*

(e) In the relaxed state of the fibrillar protoplasm, the part of the end-cell lumen which it controls must still be patent.

With these new specifications, we are in a position to revise Dahlgren's account of the operation of the end-cell. This will now require four successive stages, as follows.

First: Contraction of the fibrillar cytoplasm to enlarge the lumen of the end-cell and draw air in from the terminal tracheal twig.

Second: Contraction of the rounded mass to close off the end-cell lumen proximal to the point enlarged by the contraction of the fibrillar protoplasm.

Third: Relaxation of the fibrillar cytoplasm, so that the elastic recoil of the wall of the end-cell lumen applies pressure to the contained air and forces some of it distally into the tracheoles.

Fourth: Relaxation of the rounded mass, opening its lumen and releasing the pressure on the air in the tracheoles. Since the entire passage through the end-cell is now open, diffusion can replenish the oxygen in the tracheolar air.

Since the anatomical and operational features just outlined are so largely speculative, it is desirable to ascertain how well they meet criticisms which have been made of the concept of a mechanical end-cell valve, and how well they explain various known facts about the control mechanism.

Maloeuf raised several objections, based on his understanding that Dahlgren's theory provided for *stopping* the flash by contraction of the end-cell. This is another of the points lacking in Dahlgren's exposition. According to the revised mechanism outlined above, the decay of luminescence would be caused simply by the decrease in pO_2 due to respiration and luminescence, to a level insufficient to support luminescence. The flash is actually *started,* rather than stopped, by end-cell contraction.

---

* If such intricacy seems difficult to visualize in a structure as small as an end-cell, we should recall the complexity of some of the small Protozoa, to say nothing of that of an individual cilium.

Another obvious question to be asked about the mechanical end-cell control theory is whether the amount of air which could be forced into the photogenic tissue in a single "jet" would be sufficient to support the necessary luminescence. We are seriously hampered here by not knowing what that requirement is, but, if the demands of light-production are anything like those of ordinary tissue respiration, it seems unlikely that the mere contraction of the sphincter-like rounded mass (resulting in the inward movement, by a distance of a micron or two at most, of a fraction of a cubic micron of air) would be adequate. The same objection would probably apply to any theory involving simple constriction at some point in the lumen of the end-cell, even allowing for the possibility that the end-cell at the other end of the tracheole might provide a similar amount of air simultaneously. Under such circumstances, it seems that a preliminary enlargement of the end-cell passage, such as Dahlgren postulated, would be necessary. On the other hand, if the oxygen requirement of luminescence were very low, it might be possible to simplify the contractile mechanism to the extent of eliminating the expansion of the tracheolar lumen caused by contraction of the fibrillar protoplasm. The mechanism would then consist simply of two sphincters in tandem on the end-cell lumen, contracting and relaxing with the usual peristaltic rhythm.

Dahlgren describes the tracheoles as intercellular, and as connecting two end-cells of contiguous cylinders (anastomosis), but he does not consider specifically whether the "jet" is to be a *flow* of air through the open tube or a raising of the air pressure inside the tracheole. The former would result if the end-cell at one end of the tracheole were relaxed while the other was contracted; the latter, if both end-cells contracted at once. The former, or "flow" hypothesis, would have the following consequences:

(a) It might result in a more efficient renewal of the tracheolar oxygen, since air low in oxygen would be exhausted into a large space (the twig) with each pulse, where diffusion could raise its oxygen content quickly.

(b) It would require a high order of nervous coordination (alternate firing of the members of end-cell or cylinder pairs).

(c) Luminescence should spread from one end-cell toward its partner.

The "pressure" hypothesis would have the following consequences:

(a) It could operate with simultaneous firing of end-cells.

(b) Action would probably be faster than with a "flow" mechanism, since, in the latter, each flash would have to progress across the photogenic cell.

(c) It might result in a higher momentary oxygen tension in the tracheole than could be obtained by flow alone.

(d) Increase in oxygen tension would be simultaneous along the whole tracheole.

It seems to me that the points enumerated illustrate certain real advantages of a pressure mechanism. It has been pointed out previously that the final step in any oxygen control mechanism must be the passage of oxygen into the photogenic cytoplasm, by diffusion. Since the rate of diffusion cannot be changed, the problem of bringing about very rapid control becomes one of making diffusion distances as short as possible. Such a requirement is obviously not met by any arrangement whereby oxygen reaches the surface of the photogenic cell by diffusing down the tracheole. Nor is it met by a flow mechanism, because time is lost in raising the oxygen tension at successive points along the tracheole. With a pressure mechanism, however, most of the necessary oxygen can already be present in the tracheole, though ineffective in causing luminescence, due to its partial pressure being sub-threshold or because all is being used in respiration (see discussion of competition hypothesis). The pressure increase, therefore, results in a simultaneous rise in oxygen tension at all points in the tracheole, at a rate limited only by the speed of contraction of the end-cell. The rate of build-up of the flash is, then, determined by the speed of gaseous diffusion over a maximum distance of the radius of the tracheole, plus liquid diffusion in the photogenic cytoplasm. In the same way, the decay phase, or shutting-off of the flash, can occur as rapidly as the end-cell can relax,* because the oxygen tension falls simultaneously to a sub-threshold value at every point in the tracheole.

Another, if less crucial, advantage of the pressure hypothesis is that it would also apply if an inner section of the tracheole were blocked or restricted by liquid or tissue. The only effect of liquid in part of the tracheole would be to reduce the intensity of the flash by impeding oxygen access to part of the photogenic cytoplasm. In fact, such fluid might even be assigned a possible function in so varying flash intensity, thus taking care of those reports which describe fluid in the tracheoles. Moreover, since fluid would not affect the duration of the flash, one explanation is provided for Snell's observation that, although the intensity of normal flashes is highly variable, the duration is astonishingly constant. It is thus one of the advantages of the mechanical end-cell theory, in the revised form presented here, that it would operate whether the tracheoles are extracellular or intracellular, or whether they do or do not anastomose (provided that, if they do, the two end-cells of a tracheole contract simultaneously).

This whole superstructure is so complex and highly theoretical that it may seem incongruous to speak of considering the "evidence" for one or the other hypothesis. Nevertheless, it must be kept in mind that the control of luminescence in the firefly, no matter how it is brought about, is, after all, a wonderfully precise accomplishment. It would thus be surprising, indeed, if its mechanism were not intricate, both anatomically and physiologically. At any rate, the observations of Lund, Alexander, and others, that luminescence spreads out from the cylinder wall, might be thought to favor the "flow" hypothesis. However, it must be remembered that all such observations were of glows rather than flashes and, as we have seen, the glow is most reasonably explained as due merely to diffusion through an inactivated control mechanism. In terms of the present theory, this "inactivation" would mean the relaxation (opening) of the rounded mass, without any activity of the fibrillar cytoplasm.

---

* Since relaxation decreases pO_2 in the tracheole below the threshold for luminescence, decay is determined only by depletion of the oxygen actually in the cytoplasm. According to the revised version of the mechanical end-cell theory, the "contraction", or action which forces air into the tracheole would actually be a relaxation of the fibrillar cytoplasm, the force being provided by elastic recoil of the wall of the end-cell lumen. The "relaxation" would be the relaxation of the "rounded body."

Further important evidence on the mechanical theory of end-cell action can be derived from considering whether it can also explain the phenomena discussed in connection with glowing. Alexander strongly emphasized the distinction between glow control and flash control, and apparently thought of the two as being quite separate, although most of his conclusions on flash control were derived from observations of glows. The points which particularly seem to indicate a dual control are the following:

(a) Voluntary flashing is abolished in the dead or decapitated animal, whereas glowing is often brilliant.

(b) Normal or electrically stimulated flashes usually occur independently of glowing.

(c) Voluntary flashing is abolished during glows produced by anesthetics or by high or low oxygen. During the latter, at least, flashing cannot even be elicited by electrical stimulation.

(d) No glow approaches in intensity the maximum reached in the flash.

(e) Under some circumstances, flashes can occur superimposed on a glow (early stages of high oxygen effect—Alexander; early stages of anesthesia or water injection—Buck).

(f) The decay portion of the pseudoflash time-intensity curve is far more gradual than that of the normal flash (Snell, Alexander).

All these observations can be explained as follows, on the basis of a valvular end-cell which has the properties which I have postulated:

(a) In the normal non-luminescent condition, the rounded body is in a state of tonic contraction such that its lumen admits only enough air to supply respiration, and the fibrillar protoplasm is relaxed, so that the part of the lumen it controls is minimal.

(b) In response to voluntary or to applied electrical stimulus, there occurs the four-stage contraction-relaxation sequence postulated as producing the air "jet." At the end of the resultant flash, the rounded body returns to its original state, with an opening insufficient to permit glow. (Flash without glow.)

(c) Anesthetics or death relax the rounded body enough so that sufficient air diffuses through to exceed the oxygen threshold for glowing. Or, alternatively, these factors depress respiration, so that oxygen becomes available for photogeny. They also inactivate the fibrillar cytoplasm, which is already relaxed (open), so that no flash is possible. (Glow without flash.)

(d) The assumption that, in ordinary air at one atmosphere pressure, glow intensity is determined only by the aperture through the rounded body, would explain why the glow (diffusion-controlled) is not as intense as the flash (pressure-controlled).

(e) A slight stimulation, preceding anesthetic or toxic effect, is a well-known pharmacological phenomenon. A slight stimulation of the fibrillar protoplasm of the end-cell, at the same time that the rounded body reached its relaxed phase, would well explain flashes superimposed on glows.*

(f) The slow decay of the pseudoflash, as compared with that of the normal flash, can be explained by a slowed-down response of the rounded body as it recovers from anoxia and resumes its tonic state of partial contraction.

In summary, the facts and assumptions presented in the last few pages suffice to explain many of the known facts about light-emission in the firefly, on the basis of mechanical control of oxygen by the end-cell. There is, however, no strong evidence that the concept is more than a hypothesis, except the correlation between the possession of end-cells and the ability to flash. Moreover, there is no proof that oxygen limitation operates directly to control luminescence. On the other hand, there exists also the hypothesis of direct nerve action, which explains known phenomena as satisfactorily as the end-cell theory. This hypothesis is supported by the fact that intensity of luminescence varies to some extent with intensity of stimulation in forms without end-cells, and by a variety of other suggestive but not "air-tight" experiments, particularly those of Bellesme, Steinach, and Snell. An obvious question is whether there is any possibility of reconciling the two ideas, so that the presence of end-cells is explicable while the apparent physiological advantages of direct nerve action are retained.

---

* As an additional refinement, it could be assumed that the rounded body may also contract (close) initially in response to a brief stimulation by anesthetics. This would explain Gerretsen's observation of an initial transitory suppression of luminescence by anesthetics, which Maloeuf used as an argument against the mechanical control of flashing.

An ostensible reconciliation was achieved in Brücke's (1881) suggestion that oxygen exercises control indirectly, no luminescence occurring when the direct nerves to the photogenic tissue are kept anaerobic. There is little anatomical support for this idea, nor is it easy to reconcile with, *e.g.*, the hypoxic glow. Furthermore, it apparently requires a second set of "indirect" nerves for regulating the oxygen (*e.g.*, *via* end-cells). However, the possibility of reconciling the two major hypotheses of control is so alluring that it is to be hoped that some further evidence will be adduced which points to an "indirect" function of the end-cell.

On the whole, the strongest argument against direct control of luminescence by oxygen is the fact that no other similar mechanism is known. It is true that rapid metabolic responses are, as yet, very imperfectly understood, but in the few instances which have been at least partially worked out (nerve, muscle, activation of dormant respiration), all the evidence points to enzymatic activity as the physiological "trigger" which sets off the reaction. Such enzymatic activity, in other words, appears to be the usual method of bringing about metabolic changes which are sudden and complete. There is no point in attempting to propose any specific mechanism in regard to control of luminescence, while the analogous control of muscular contraction is still unclear after an immense amount of good experimental work. However, it can be anticipated that the study of enzymatic inhibition and activation in the luminescent system will prove to be a fruitful field in future research on the control problem.

SUMMARY

In spite of the many morphological and physiological data which concern luminescence in the firefly, there seem to be surprisingly few unequivocal major conclusions which can be drawn. This is due partly to the many differences which exist between various species of fireflies and which often make generalizations impossible, and partly to the fact that our knowledge is seriously deficient in many essential points. Among the generalizations which can be made are the following:

1. The luminous organs of fireflies vary greatly in size, shape and position, and with sex and developmental stage.

2. Firefly light organs can be divided into six histological types: (1) those with no specific tracheal supply; (2) those in which the tracheae show tree-like branching and in which a "reflector" layer internal to the actual photogenic tissue is lacking; (3) those like Type 2 except for having a "reflector" tissue; (4) those in which the tracheae run through the reflector layer, branch out in the interface between reflector and photogenic layers, and terminate in "tracheal end-cells" from which minute tracheal capillaries or tracheoles run into the photogenic tissue; (5) those like Type 4 except that the tracheae branch within the photogenic layer before terminating in end-cells and tracheoles; (6) those in which the tracheae run vertically through the photogenic layer in tissue rods called "cylinders", which contain the tracheal twigs and usually end-cells, and from which the tracheoles pass into the photogenic tissue. This sixth type occurs in most of the common adult American fireflies, while the third type is characteristic of firefly larvae.

3. The reflector layer has been shown to differ chemically and morphologically from the photogenic layer. No clear-cut evidence has been adduced as to the function or functions of the reflector layer, although numerous writers have postulated metabolic connections between it and the photogenic layer.

4. The location and morphology of the tracheal end-cell, and the fact that it is the chief site of reduction of inspired osmium tetroxide vapor, have engendered many suggestions that it functions in controlling the oxygen used in luminescence. However, tracheal end-cells, of which there appear to be at least two types, show no conclusive morphological evidence of being able to function in the way postulated.

5. Microscopic observation has shown that the light originates in the photogenic cells. The granules which are a characteristic feature of these cells are almost universally regarded as being the source of light, but no morphological evidence exists which confirms this assumption, nor the view, held by some investigators, that the granules are symbiotic luminous bacteria.

6. The anatomical course of firefly tracheoles strongly suggests that they function in conducting air to the photogenic tissue, and there is strong inferential evidence of the normal presence of air in the tracheoles, although this has not been demonstrated directly. Great variation exists in different organs, but in the most complex type the majority of evidence indicates that the tracheoles pursue an intercellular, rather than intracellular, course and anastomose directly with tracheoles from adjacent cylinders.

7. Nerves to photogenic cells or end-cells, or both, have been described. Experimental evidence indicates strongly that the nervous system, while not necessary for the production of light, does play an essential role in the control of luminescence in the normal living firefly. No conclusion can yet be drawn as to whether the nervous control is exerted directly upon the photogenic tissue, or secondarily (by way of the oxygen supply, for example). However, there is considerable evidence suggestive of the former.

8. By silver nitrate impregnation, a structure can be demonstrated which resembles an ultra-tracheolar network of tubules connecting neighboring tracheoles.

9. The normal types of light-emission in various fireflies fall into four increasingly complex classes: (a) the continuous glow; (b) the intermittent glow; (c) the pulsation; (d) the flash (of which there are several types). If various kinds of fireflies are arranged in order of increasing complexity of the air-supply to the photogenic organ, that order corresponds well to the above sequence of complexity of light-emission.

10. In the more complex types of firefly luminous organ, there are many thousands of photogenic units, each of which is active for a few hundredths or perhaps thousandths of a second.

11. Several hypothetical mechanisms for intracellular control of luminescence have been proposed.

12. It is probable that neither water, luciferin, nor luciferase is a limited reactant in the normal control of luminescence in the firefly.

13. It is certain that oxygen is essential for continued luminescence. and possible that it is the regulated reactant in the control of light-production in the firefly.

14. The spiracles have no causal connection with the control of normal flashing, or of glows caused by various experimental treatments. It is likely, however, that spiracular opening is prerequisite to long-continued luminescence.

15. The "internal control" of luminescence responds in a graded, rather than in an all-or-none fashion, and over-compensates in "opening" in response to lowered oxygen.

16. The simplest assumption in regard to the normal states of both spiracles and internal control is that they are in tonic contraction (closed). Partial anoxia, anesthesia and similar treatments would, thus, produce a relaxed (open) condition, permitting luminescence.

17. There is no anatomical evidence specifically identifying the internal control with the tracheal end-cell, although in the more complex types of firefly organ, where luminescence can be controlled rapidly, there is strong indirect evidence that the tracheal end-cell is concerned in control. However, firefly larvae, which lack end-cells, exhibit internal control of glowing similar to that in adults, indicating that end-cells need not be concerned with glow control even if they are concerned in flash control.

18. Some theoretical implications of the hypothesis that light-production is controlled by competition for oxygen between the luminescent reaction and respiration are discussed.

19. The hypothesis that access of oxygen to the luminescent tissue is controlled by changes in the fluid content of the tracheoles, is discussed and found to be deficient in some respects.

20. The hypothesis that oxygen access is limited by a valvular action of the tracheal end-cells is discussed in detail and found to accord well with what is actually known at the moment about control of flashing in the firefly. Like others, however, the end-cell hypothesis requires the use of a number of assumptions which, so far, it has been impossible to test experimentally.

21. Arguments are presented for regarding enzymatic mechanisms as better fitted than oxygen regulation for the rapid and complete control of luminescence.

Almost all of the work reported in this paper was concerned, either directly or indirectly, with the general question of what mechanisms, external to the photogenic cell, affect its luminescence. In considering the possible trends of future work in this field, it seems clear that important new departures cannot be expected until definite solutions are obtained to a number of long-recognized basic problems. Most of these problems center around the question of whether the photogenic cell is stimulated by direct nerve action, presumably in an environment always adequate in oxygen, or whether the stimulus which sets off luminescence is the passage of oxygen into the cell.

As already related, it has so far been impossible to devise a physiological experiment by which the respective effects of nerve action and of oxygen can be separated with certainty. One of the most promising future experiments is the simultaneous recording of light-emission and of the spontaneous action potentials of the central nervous system. Yet, even if a good correlation were found between the potential pattern and luminosity, direct nervous action would not be proved. It could be argued that the nerves had been stimulating some indirect agency, such as the end-cells, which, in turn, brought about direct stimulation of the photogenic cell by oxygen. Another significant type of approach involves a careful physiological study of two species of lampyrid fireflies, one of which has end-cells and the other has not.

The most direct lines of attack, at the moment, seem anatomical. Conclusive establishment of the ultimate terminations of the nerves to the photogenic organ, if carried out both in fireflies with end-cells and in those without, might go far in resolving the present impasse. Conversely, any cytological evidence bearing on the structure and activity of the end-cells and tracheoles in the living photogenic organ, both in the luminescent and non-luminescent state, could hardly fail to be significant. Likewise, evidence that end-cells can regulate oxygen rapidly in non-luminous insects, would be of great value. Furthermore, it would be of at least theoretical interest to calculate the potentialities of gaseous diffusion, flow, and pressure, in supplying air to the photogenic tissue through conducting systems of the known anatomical dimensions.

Further important information on the possibility of direct nerve action might be obtained by carefully controlled experiments on electrical stimulation of the nerve cord, with special reference to the relation between length or strength of stimulus and total resulting luminescence, total luminescence in relation to oxygen tension, and total luminescence in relation to length of preceding anaerobiosis. Oxygen tensions used should be kept within the range of 100 to 300 mm., in which many physiological reactions remain constant, and particular attention should be paid to fireflies lacking end-cells.

Finally, the simplest and most important desideratum in future work on fireflies is improvement in the design of experiment. It cannot fail to impress anyone familiar with the literature, that an enormous amount of work has been weakened or negated by failure to observe such elementary methodological precautions as the use of adequate controls, frequent repetitions, the variation of only one experimental factor at a time, and appreciation of the range and frequency of the normal and experimental variation of the material.

BIBLIOGRAPHY*

Alexander, Robert S.
1943. Factors controlling firefly luminescence. J. Cell. & Comp. Physiol. **22**: 51–71.

Allard, H. A.
1931. The photoperiodism of the firefly *Photinus pyralis* Linn. Its relation to the evening twilight and other conditions. Proc. Ent. Soc. Wash. **33**: 49–58.

Arnold, Carl
1881. Beiträge zur vergleichenden Physiologie. Mittheil. Naturforsch. Gesellsch. Bern 1880: 151-192, esp. 175-178.

Ball, Eric, & Pauline A. Ramsdell
1944. The flavin-adenine dinucleotide content of firefly lanterns. J. Am. Chem. Soc. **66**: 1419.

Barber, H. S.
1941. Species of fireflies in Jamaica (Coleoptera, Lampyridae). Proc. Rochester Acad. Sci. **8**: 1–13.

* No attempt has been made to present a complete bibliography on fireflies, but I have tried to include all the important contributions, particularly those which have appeared since Dahlgren's review. Wherever possible, only a single paper (the most comprehensive, rather than the first) is cited for a given author. No references are included which are not mentioned in the text. References marked with an asterisk were not available in the original, and are quoted from the reviews of **Edwards** (1868), **Kerville** (1881, 1887), **Henneguy** (1904), **Höllrigl** (1908), **Berlese** (1909), **Mangold** (1910), **Kastle & McDermott** (1910), **Harvey** (1920, 1940, and 1941), and **Pratje** (1923). I am indebted to Professor Harvey for the opportunity to read the papers of Hasama, which were not otherwise available in this country.

Bellesme, Jousset de
1880. Recherches expérimentales sur la phosphorescence du lampyre. J. Anat. & Physiol. **16:** 121–169.

Berlese, Antonia
1909. Gli Insetti **I:** 709-714. Societa Editrice Libraria. Milano.

Bongardt, Johannes
1903. Beiträge zur Kenntnis der Leuchtorgane einheimischer Lampyriden. Z. wiss. Zool. **75:** 1–44.

Brown, Dugald E. S., & Cecil V. King
1931. The nature of the photogenic response of *Photuris pennsylvanica.* Physiol. Zool. **4:** 287-293.

Brücke, Ernst
1881. Vorlesungen über Physiologie **I:** 59-61. III. Aufl. Braumüller. Wien.

Buchner, Paul
1914. Sind die Leuchtorgane Pilzorgane? Zool. Anz. **45:** 17-21.
1930. Tier und Pflanze in Symbiose: 729-732. 2d Ed. Borntraeger. Berlin.

Buck, John B.
1937a. Studies on the firefly. I. The effects of light and other agents on flashing in *Photinus pyralis,* with special reference to periodicity and diurnal rhythm. Physiol. Zool. **10:** 45-58.
1937b. Studies on the firefly. II. The signal system and color vision in *Photinus pyralis.* Physiol. Zool. **10:** 412-419.
1938. Synchronous rhythmic flashing of fireflies. Quart. Rev. Biol. **13:** 301-314.
1940. Comparative histology of Coleopteran photogenic organs. Anat. Rec. **78:** suppl. 176.
1941. Studies on the firefly. III. Spectrometric data on thirteen Jamaican species. Proc. Rochester Acad. Sci. **8:** 14-21.
1942. Problems in the distribution and light organ structure of Jamaican lampyrid fireflies. Yearbook Am. Phil. Soc.: 124-129.
1946a. Some aspects of the histology and physiology of luminescence in "railroad worms." Biol. Bull. **91:** 226. (Full report in preparation.)
1946b. The spiracular factor in the control of luminescence in the firefly. Anat. Rec. **96:** 51. (Full report in preparation.)

Bugnion, E.
1929. Le ver-luisant provençal et la luciole niçoise. Mém. Assoc. des Naturalistes de Nice et des Alpes-Maritimes. (Suppl. *Riviera Scientifique* 1929.)

Bult, Tamme
1939. Over de beweging der vloeistof in de tracheolen der insecten. Proefschrift. Rijks-Universiteit te Groningen.

Burge, W. E.
1916. Comparison of the intensity of oxidation in luminous and non-luminous insects. J. Franklin Inst. **182:** 263-264.

Carradori, Joachim
1797. Objections contre l'opinion du Profes^r Spallanzani, sur la cause du luisant des Phosphores naturels. Ann. Chim. **24:** 216-225.

*** Carrara, M.**
1836. Sulla phosphorenza della lucciole comune (*Lampyris italica* L.) Biblioth. Ital. **82:** 357-370.

Carus
1864. Expériences sur la matière phosphorescente de la *Lampyris italica;* action de l'eau pour rendre à la matière desséchée cette phosphorescence. C. R. Acad. Sci. **59:** 607-608.

Coblentz. W. W.
1912. A Physical Study of the Firefly. Carnegie Inst. Washington **164.**

Creighton, William S.
1926. The effect of adrenalin on the luminescence of fireflies. Science **63**: 600-601.

Dahlgren, Ulric
1917. The production of light by animals. J. Franklin Inst. **183**: 79-94, 211-220, 323-348, 593-624.

Dahlgren, U., & W. A. Kepner
1908. A Textbook of the Principles of Animal Histology. Macmillan. New York.

Davy, H.
1810. Phil. Trans. Roy Soc. **100**: 287. (In **Macartney**.)

Dubois, Raphael
1886. Contributions à l'étude de la production de la lumière par les êtres vivants. Les Elatérides lumineux. Bull. Soc. Zool. France **11**: 1-275.

Edwards, H. Milne
1863. Leçons sur la Physiologie et l'Anatomie Comparée de l'Homme et des Animaux **VIII**: 95-105. Masson. Paris.

Eimer, Th.
1872. Bemerkungen über die Leuchtorgane von *Lampyris splendidula*. Arch. mikrosk. Anat. **8**: 652-653.

Emerson, George A.
1935. Some effects of ether on bioluminescence in the lampyrid, *Photuris pennsylvanica*. Proc. Soc. Exp. Biol. & Med. **33**: 36-40.

Emerson, George A., & Marjorie J. Emerson
1941. Mechanism of the effect of epinephrine on bioluminescence of the firefly. Proc. Soc. Biol. & Med. **48**: 700-703.

Emery, Carlo
1884. Untersuchungen über *Luciola italica* L. Z. wiss. Zool. **40**: 338-355.

Faraday, M.
1814. Journal of 1814. In: **Bence Jones.** Life and letters of Faraday 1870. **1**: 141-146. Longmans Green. London.

Forster, G.
1782. Ein Versuch mit dephlogistisirter Luft. Göttingisches Mag. Wissensch. & Lit. **3**(2): 281-288.

Fuchs, Sigmund
1891. Einige Versuche an den Leuchtorganen von *Elater noctilucus* L. Zentralbl. Physiol. **5**: 321-325.

Geipel, Erich
1915. Beiträge zur Anatomie der Leuchtorgane tropischer Käfer. Z. wiss. Zool. **112**: 239-290.

Gerretsen, F. C.
1922. Einige Notizen über das Leuchten des javanischen Leuchtkäfers (*Luciola Vittata* Cast.) Biol. Zentralbl. **42**: 1-9.

Grinfeld, Rafael
1944. Contribucion al estudio del espectro de la luz de las luciernagas. Contr. Facultád de Ciencias Físicomatemáticas La Plata **3**: 447-461.

Grotthuss, Theodore de
1807. Sur la combinaison du phosphore avec les métaux et leurs oxides par la voie humide, etc. Ann. Chim. **64**: 19-41.

Harvey, E. Newton
1920. The Nature of Animal Light. Lippincott. Philadelphia.
1922. The permeability of cells for oxygen and its significance for the theory of stimulation. J. Gen. Physiol. **5**: 215-222.
1931. Photocell analysis of the light of the Cuban elaterid beetle, *Pyrophorus*. J. Gen. Physiol. **15**: 139-145.
1940. Living Light. Princeton University Press.
1941. Review of bioluminescence. Ann. Rev. Biochem. **10**: 531-552.

1944. The nature of the red and green luminescence of the South American "railroad worm," *Phryxothrix*. J. Cell. & Comp. Physiol. **23**: 31-38.
1945. Note on the red luminescence and the red pigment of the "railroad worm." J. Cell. & Comp. Physiol. **26**: 185-187.

Harvey, E. Newton, & Robert T. Hall
1929. Will the adult firefly luminesce if its larval organs are entirely removed? Science **69**: 253-254.

Hasama, Bun-ichi
1942a. Über die Biolumineszenz bei *Pyrocoelia rufa* im Aktionsstrombild sowie im histologischen Bild. Annot. Zool. Jap. **21**: 59-77.
1942b. Über die Biolumineszenz der *Luciola lateralis* im zytologischen Bild sowie im Potentialbild ihres Leuchtorgans. Cytologia **12**: 366-377.
1942c. Über die Biolumineszenz der Larve von *Luciola cruciata* sowie von *Pyrocoelia rufa* im Aktionsstrombild und im histologischen Bild ihres Leuchtogans. Cytologia **12**: 378-388.
1942d. Zytologische Untersuchung des Leuchtorgans von *Luciola cruciata*. Cytologia **12**: 389-396.
1942e. Zytologische Untersuchungen des Leuchtorgans von zwei tropischen Leuchtkäfern, *Pyrocoelia analis* und *Luciola Gorhami*. Cytologia **12**: 486-494.
1944a. Entwicklung des imaginalen Leuchtorgans der *Luciola cruciata* in histologischer sowie bioelektrischer Hinsicht. Cytologia **13**: 155-161.
1944b. Histologische Untersuchungen des Leuchtorgans der *Luciola parvula*. Cytologia **13**: 179-185.

Heinemann, Carl
1872. Untersuchungen über die Leuchtorgane der bei Vera-Cruz vorkommenden Leuchtkäfer. 1. Abtheil. Arch. Mikr. Anat. **8**: 461-471.
1873. Aschenanalyse von Leuchtorganen mexikanischer Cucujos. Arch. ges. Physiol. **7**: 365-366.
1886. Zur Anatomie und Physiologie der Leuchtorgane mexikanischer Cucujos (*Pyrophorus*). Arch. Mikr. Anat. **27**: 296-382.

Heller, Joh. Flor.
1853. Ueber das Leuchten im Pflanzen- und Thierreiche. Arch. physiol. & pathol. Chemie & Mikroskopie **6**: 44-54; 81-90; 121-137; 161-166; 201-216; 241-251; esp. 203-209.

Henneguy, L. Felix
1904. Les Insectes: 92-97. Masson. Paris.

Hermbstädt
1808. Bemerkungen über das Leuchten organischer Körper im Leben und nach dem Tode derselben. Ges. Naturf. Freunde, Mag. neuesten Entdeck. Naturk. **2**: 248-256.

Hess, W. N.
1920. Notes on the biology of some common Lampyridae. Biol. Bull. **38**: 39-76.
1921. Tracheation of the light-organs of some common Lampyridae. Anat. Rec. **20**: 155-161.
1922. Origin and development of the light-organs of *Photuris pennsylvanica* De Geer. J. Morph. **36**: 245-277.

Höllrigl, M. Gregoria
1908. Lebensgeschichte von *Lamprorrhiza splendidula* mit besonderer Berücksichtigung des Leuchtvermögens. Ber. naturwiss.—med. Verein Innsbruck **31**: 169-230.

Holmgren, Emil
1895. Die trachealen Endverzweigungen bei den Spinndrüsen der Lepidopterenlarven. Anat. Anz. **11**: 340-346.

Hoskins, W. M.
1940. Recent contributions of insect physiology to insect toxicology and control. Hilgardia **13**: 307-386.

Johnson, F. H., H. Eyring, R. Steblay, H. Chaplin, C. Huber, & G. Gherardi
1945. The nature and control of reactions in bioluminescence with special reference to the mechanism of reversible and irreversible inhibitions by hydrogen and hydroxyl ions, temperature, pressure, alcohol, urethane, and sulfanilamide in bacteria. J. Gen. Physiol. **28**: 463-537.

Joseph, Gustav
1854. Beobachtungen über das Leuchten der Johanniskäfer. Z. Entom. **8** (Coleoptera): 1-12.

Kastle, Joseph H., & F. Alex. McDermott
1910. Some observations on the production of light by the firefly. Am. J. Physiol. **27**: 122-151.

Kerville, Henri Gadeau de
1881. Les Insectes Phosphorescents. Leon Deshays. Rouen.
1887. Les Insectes Phosphorescents. Notes Complémentaires et Bibliographie Générale. Julien Lecerf. Rouen.

Knoche
1910. Unpublished experiments described by **Mangold** (p. 357).

Kölliker, A. v.
1858. Die Leuchtorgane von *Lampyris,* eine vorläufige Mittheilung. Verh. phys. med. Ges. Würzburg **8**: 217-224.

Krogh, August
1919. The rate of diffusion of gases through animal tissues, with some remarks on the coefficient of invasion. J. Physiol. **52**: 391-408.

Kuhnt, P.
1907. Das Leuchten der Lampyriden. Entomol. Wochenblatt **24**: 3-4.

Leydig, Franz
1857. Lehrbuch der Histologie des Menschen und der Thiere: 342-344. Meidinger Sohn & Co. Frankfurt a. M.

Lindemann, Carl
1863. Anatomische Untersuchungen über die Struktur des Leuchtorganes von *Lampyris splendidula.* Bull. Soc. Imp. Naturalistes Moscou **36**(2): 437-456.

Lund, E. J.
1911. On the structure, physiology and use of photogenic organs, with special reference to the Lampyridae. J. Exp. Zool. **11**: 415-467.

Macaire, J.
1821. Mémoire sur la phosphorescence des lampyres. J. Phys., Chim., Hist. Nat. & Arts **93**: 46-56. Repeated in Ann. Chim. & Phys. **17**: 151-167 (misprinted as 251-267).

Macartney, J.
1810. Observations on luminous animals. Phil. Trans. Roy. Soc. **100**: 258-293.

Maloeuf, N. S. R.
1938. The basis of the rhythmic flashing of the firefly. Ann. Ent. Soc. Am. **31**: 374-380.

Mangold, E.
1910. Die Produktion von Licht. Hans Winterstein's Handbuch der vergleichenden Physiologie **3**(2): 225-392.

Mast, S. O.
1912. Behavior of fire-flies (*Photinus pyralis*)? with special reference to the problem of orientation. J. Animal Behavior **2**: 256-272.

Matteucci, C.
1843. Sur la phosphorescence du lampyre d'Italie (*L. italica*). C. R. Acad. Sci. **17**: 309-312.

McDermott, F. A.
1910. A note on the light-emission of some American Lampyridae. Canad. Entomol. **42**: 357-363.

1911a. Some further observations on the light-emission of American Lampyridae: the photogenic function as a mating adaptation in the photinini. Canad. Entomol. **43**: 399-406.

1911b. The stability of the photogenic material of the Lampyridae and its probable chemical nature. J. Am. Chem. Soc. **33**: 1791-1796.

1912. Observations on the light-emission of American Lampyridae. Fourth paper. Canad. Entomol. **44**: 309-311.

1914. The ecologic relations of the photogenic function among insects. Z. wiss. Insektenbiologie **10**: 303-307.

1915. Experiments on the nature of the photogenic processes in the Lampyridae. J. Am. Chem. Soc. **37**: 401-404.

1917. Observations on the light-emission of American Lampyridae: The photogenic function as a mating adaptation; fifth paper. Canad. Entomol. **49**: 53-61.

McDermott, F. A., & C. G. Crane

1911. A comparative study of the structure of the photogenic organs of certain American Lampyridae. Am. Nat. **45**: 306-313.

McElroy, William D., & Robert Ballentine

1944. The mechanism of bioluminescence. Proc. Nat. Acad. Sci. **30**: 377-382.

Müller, Philipp Wilbrand Jakob

1805. Beiträge zur Naturgeschichte des halbdeckigen Leuchtkäfers, *Lampyris hemiptera* Fabr. Illiger's Mag. Insektenk. **4**: 175-196.

Okada, Yo K.

1935a. Origin and development of the photogenic organs of lampyrids, with special reference to those of *Luciola cruciata* Motschulsky and *Pyrocoelia rufa* Ern. Olivier. Mem. Coll. Sci., Kyoto Imp. Univ. Ser. B **10**: 209-228.

1935b. Luminous apparatus in lampyrids. Shokubutsu oyohi Dobutsu ("Botany and Zoology") **3**: 1312-1318, 1475-1482, 1638-1648, 1799-1806. (Unfortunately, I have not yet been able to have this extensive and beautifully illustrated paper translated. Citations are from figure headings, part of which are in English.)

Osten-Sacken, Baron v.

1861. Die amerikanischen Leuchtkäfer. Stettiner Ent. Ztg. **22**: 54-55.

Owsjannikow, Ph.

1864. Ueber das Leuchten der Larven von *Lampyris noctiluca*. Bull. Acad. Imp. Sci. St. Petersburg **7**: 55-61.

1868. Ein Beitrag zur Kenntniss der Leuchtorgane von *Lampyris noctiluca*. Mém. Acad. Imp. Sci. St. Petersburg Ser. VII. **11**(17).

Perkins, Michael

1931. Light of glow-worms. Nature **128**: 905.

Peters, Wilhelm

1841. Ueber das Leuchten der *Lampyris italica*. Arch. Anat. & Physiol.: 229-233.

Pflüger, E.

1875. Beiträge zur Lehre von der Respiration. I. Ueber die physiologische Verbrennung in den lebendigen Organismen. Arch. ges. Physiol. **10**: 251-367. (The part relevant to this paper is his fifth section: V. Die Phosphoreszenz der lebendigen Organismen und ihre Bedeutung für die Prinzipien der Respiration—275-300.

Pierantoni, U.

1914. Sulla luminosità e gli organi luminosi di *Lampyris noctiluca* L. Boll. Soc. Nat. Napoli **27**: 83-88.

Pratje, Andre

1923. Das Leuchten der Organismen: eine Übersicht über die neuere Literatur. Ergeb. Physiol. **21**: 166-273.

Prowazek, S.

1908. Unpublished work cited by Steche.

Quatrefages, A. de
1850. Mémoire sur la phosphorescence de quelques invertébrés marins. Ann. Sci. Nat. Ser. III (Zool.) **14:** 236-281.

Rau, Phil
1932. Rhythmic periodicity and synchronous flashing in the firefly, *Photinus pyralis,* with notes on *Photuris pennsylvanicus.* Ecology **13:** 7-11.

***Remy, P.**
1925. Contribution à l'étude de l'appareil respiratoire et de la respiration chez quelques invertébrés. Vagner. Nancy.

Richards, A. Glenn
1947. The organization of Arthropod cuticle: a modified interpretation. Science **105:** 170-171.

Richards, A. Glenn, Jr., & Thomas F. Anderson
1942. Electron micrographs of insect tracheae. J. N. Y. Ent. Soc. **50:** 147-167.

Robin, Ch. & A. Laboulbène
1873. Sur les organes phosphorescents thoraciques et abdominal du cocuyo de Cuba (*Pyrophorus noctilucus; Elatei noctilucus,* L.). C. R. Acad. Sci. **77:** 511-517.

Schultze, Max
1865. Zur Kenntnis der Leuchtorgane von *Lampyris splendidula.* Arch. mikrosk. Anat. **1:** 124-137.

Seaman, Wm. H.
1891. On the luminous organs of insects. Proc. Am. Soc. Microsc. **13:** 133-162.

Seifter, Joseph
1945. An unusual action of amphetamine. Science **102:** 597.

Severn, H. A.
1881. Notes on the Indian glow-fly. Nature **24:** 165.

Shafer, G. D.
1911. The effect of certain gases and insecticides upon the activity and respiration of insects. J. Ecol. Entomol. **4:** 47-50.

Shoup, Charles S.
1929. The respiration of luminous bacteria and the effect of oxygen tension upon oxygen consumption. J. Gen. Physiol. **13:** 27-45.

Siebold, C. Th. v.
1848. Lehrbuch der vergleichenden Anatomie der wirbellosen Tiere (pp. 632-633). Veit & Co. Berlin.

Snell, Peter A.
1932. The control of luminescence in the male lampyrid firefly, *Photuris pennsylvanica,* with special reference to the effect of oxygen tension on flashing. J. Cell. & Comp. Physiol. **1:** 37-51.

Snyder, Charles D., & Aleida v.'t H. Snyder
1920. The flashing interval of fireflies—its temperature coefficient—an explanation of synchronous flashing. Am. J. Physiol. **51:** 536-542.

Spallanzani, Lazzaro
1796. Chimico Esame Degli Esperimenti del Sig. Gottling, Professore a Jena, Sopra la Luce del Fosfor di Kunkel etc. Presso La Società Tipografica. Modena. (See paragraphs 114-130.)

Steche, O.
1908. Beobachtungen über das Leuchten tropischer Lampyriden. Zool. Anz. **32:** 710-712.

Steinach, E.
1908. Die Summation einzeln unwirksamer Reize als allgemeine Lebenserscheinung. Vergleichend-physiologische Untersuchungen. Arch. ges. Physiol. **125:** 239-346. The relevant part is: III. Leuchtzellen von *Lampyris.* (Sekretorische Zellen—pp. 284-289.)

Takagi, Syunzo

1934. Mitochondria in the luminous organs of *Luciola cruciata* Motschulsky. Proc. Imp. Acad. **10**: 692-694.

Todd, T. J.

1826. An inquiry into the nature of the luminous power of some of the Lampyrides, etc. Quart. J. Lit. Sci. & Arts **21**: 241-251.

Townsend, Anne B.

1904. The histology of the light organs of *Photinus marginellus*. Am. Nat. **38**: 127-151.

Tozzetti, Adolfo Targioni

1870. Sull' organo che fà lume nelle lucciole volanti d'Italia (*Luciola italica*). Bull. Soc. Ent. Ital. **2**: 177-189.

Verworn, Max

1892. Ein automatisches Centrum für die Lichtproduction bei *Luciola italica* L. Zentralbl. Physiol. **6**: 69-74.

Vogel, R.

1922. Uber die Topographie der Leuchtorgane von *Phausis splendidula* Leconte. Biol. Zentralbl. **42**: 138-140.

Vonwiller, P.

1921. Anatomische Bemerkungen über den Bau der Leuchtorgane von *Lampyris splendidula*. Festschr. Prof. Zschokke. Kober. Basel.

Watasé, S.

1895. On the physical basis of animal phosphorescence. Biol. Lect., Marine Biol. Lab. Woods Holl **1**: 101-118.

Weitlaner, Franz

1909. Etwas vom Johanniskäferchen (*Lampyris splendidula, noctiluca*). Verh. d. kais.-kön. zool.-bot. Ges. Wien **59**: 94-103.

Wielowiejski, Heinrich R. v.

1882. Studien über die Lampyriden. Z. wiss. Zool. **37**: 354-428.

Wigglesworth, V. B.

1930. A theory of tracheal respiration in insects. Proc. Roy. Soc. London B **106**: 229-250.

1931. The respiration of insects. Biol. Rev. **5**: 181-220.

1938a. The regulation of osmotic pressure and chloride concentration in the haemolymph of mosquito larvae. J. Exp. Biol. **15**: 235-247.

1938b. The absorption of fluid from the tracheal system of mosquito larvae at hatching and moulting. J. Exp. Biol. **15**: 248-254.

1939. The Principles of Insect Physiology. Dutton. New York.

Williams, F. X.

1916. Photogenic organs and embryology of lampyrids. J. Morph. **28**: 145-207.

Wistinghausen, C. v.

1890. Über Tracheenendigungen in den Sericterien der Raupen. Z. wiss. Zool. **49**: 565-582.

Wood, R. W.

1939. A firefly "spinthariscope." Science **90**: 233-234.

FIGURES 1-41

Key

C, cylinder; CM, cell membrane; CU, cuticle; DZ, differentiated zone of photogenic cytoplasm; EC, tracheal end-cell; EN, end-cell nucleus; EP, end-cell process; FB, fat body; M, muscle; N, nerve; NE, epithelial cell nucleus; O, region where osmium has reduced in the photogenic cytoplasm; P, photogenic layer or tissue; PC, photogenic cell; PN, nucleus of photogenic cell; R, reflector layer; RC, reflector cell; S, "rounded body"—a possible sphincter; T, trachea; TE, tracheole; TW, tracheal twig.

Unless otherwise specified in the figure legends, the preparations used for the photomicrographs are 10-micron sections, stained with Delafield's hematoxylin and eosin. Fixation was with Bouin's fluid, except for the preparation for FIGURE 32, for which hot Bouin's was used. Detailed techniques will be published elsewhere.

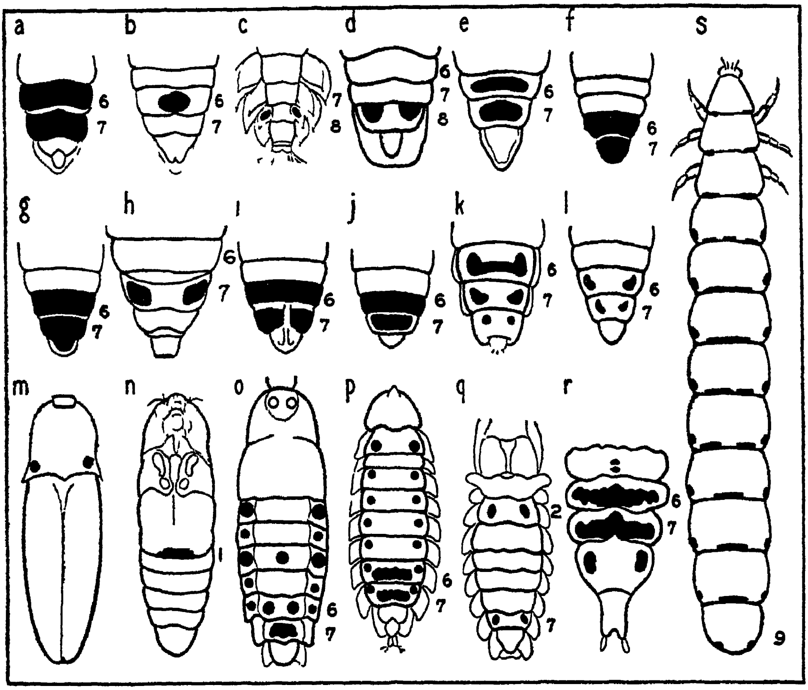


FIGURE 1. Outline diagrams showing positions and sizes of photogenic organs in representative fireflies. Photogenic organs indicated in solid black. Numbers refer to abdominal segments. Ventral views, unless otherwise noted. Sizes are variable and not shown to scale, but normal size of *Phengodes* (s) is about 40 mm., *Pyrophorus* 25 mm., entire "average" firefly 10-15 mm. a, *Photinus scintillans*, male; b, *Photinus scintillans*, female; c, *Photuris pennsylvanica*, larva; d, *Diphotus montanus*, male; e, *Photuris pennsylvanica*, female; f, *Luciola chinensis*, male; g, *Luciola cruciata*, male; h, *Luciola lusitanica*, female; i, *Luciola* sp., male; j, *Luciola lateralis*, male; k, *Pyrocoelia rufa*, female; l, *Lecontea lucifera*, female; m, *Pyrophorus noctilucus*, dorsal; n, *Pyrophorus noctilucus*, ventral; o, *Lamprorhiza* (*Lampyris*) *splendidula*, female (the number of lateral organs is quite variable, according to Vogel, 1922); p, *Phausis mulsanti*, female (the number of lateral organs is variable); q, *Phausis Delarouzeei*, nymph, dorsal; r, *Lampyris noctiluca*, female; s, *Phengodes* sp., dorsal. Figures a, b, c, and e after Hess (1920). Figures f, g, i, j, k, and l, after Okada (1935 b). Figures b, h, p, and q, after Bugnion (1929). Figure o, from various sources. Figure r, after Bongardt (1903). Figures d and s, original. Figures m and n from Seaman (1891).

FIGURES 2-7 (*see opposite page*).

FIGURE 2. Two end-cells from male of *Lamprorhiza splendidula*. Osmic acid vapor impregnation; x 650. After Bongardt (1903).

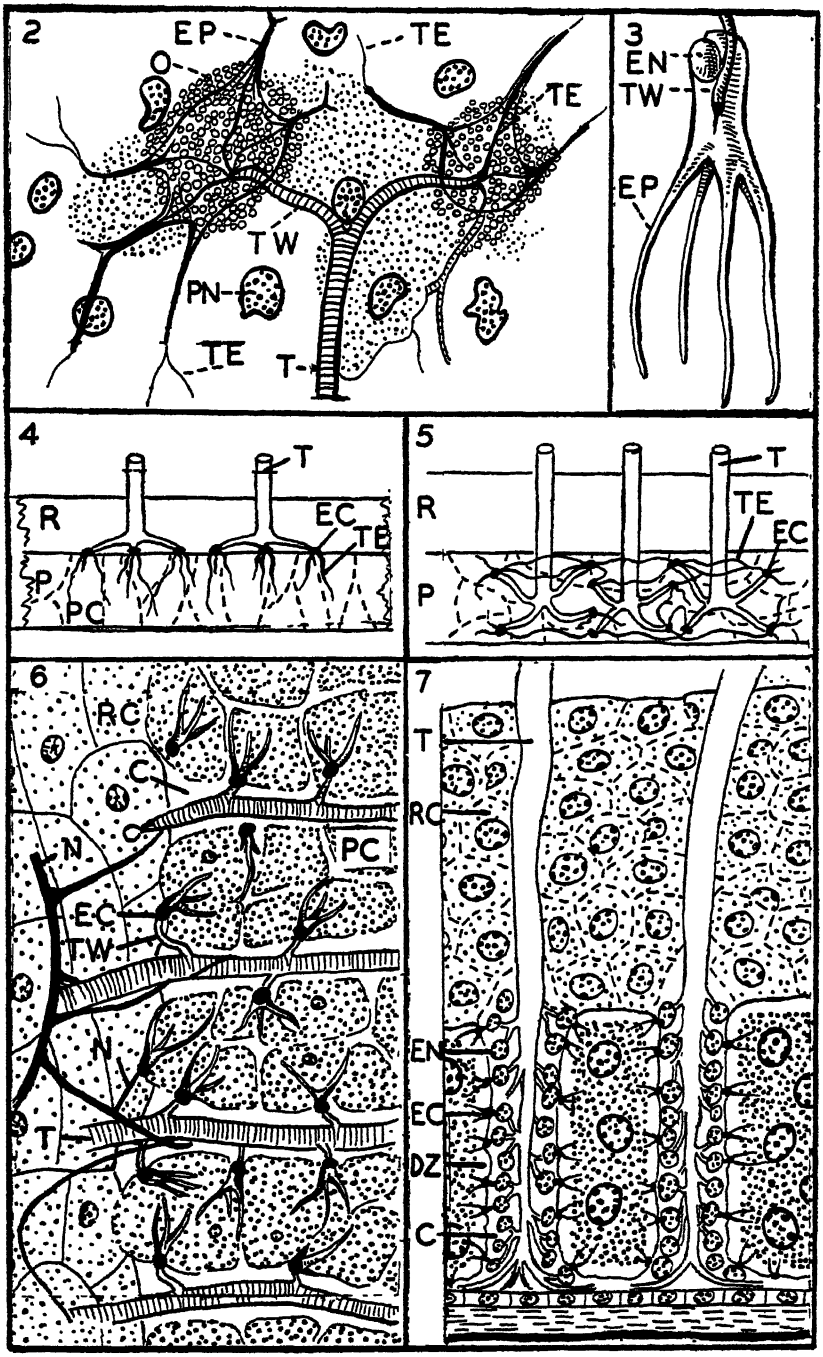
FIGURE 3. Tracheal end-cell from *Photinus marginellatus*. Osmic acid impregnation; x 1700. After Geipel (1915).

FIGURE 4. Diagram of Type 4 arrangement of tracheae. End-cells at boundary between reflector and photogenic layers, with tracheoles extending into latter. After Dahlgren (1917).

FIGURE 5. Diagram of Type 5 arrangement of tracheae. Trachea branching into interior of photogenic layer and there terminating in end-cells. After Dahlgren (1917).

FIGURE 6. Arrangement of end-cells and nerves in the light organ of *Photinus marginellatus*. After Geipel (1915).

FIGURE 7. Cross section of the light organ of adult *Photuris pennsylvanica*. After Hess (1922). Contrast with FIGURE 11.


2
EP
O
TE
TE
TW
PN
TE
T
3
EN
TW
EP
4
T
R
EC
TE
P
PC
5
T
R
TE
EC
P
6
RC
C
C
N
PC
EC
TW
N
T
7
T
RC
EN
EC
DZ
C

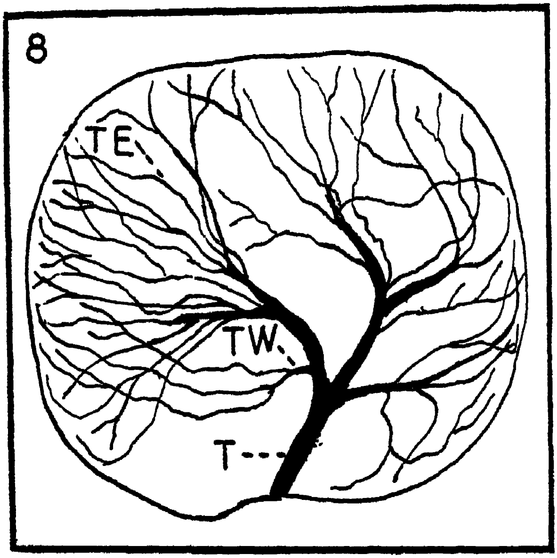


FIGURE 8. Larval organ of *Phosphaenus hemipterus* from caustic potash preparation to show tree-like branching of tracheae. After **Bongardt** (1903).

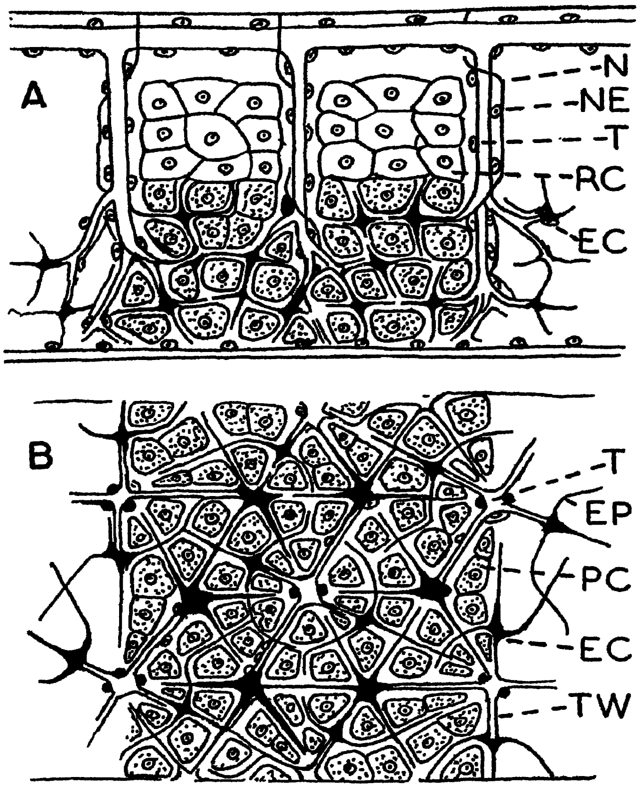


FIGURE 9. Diagram of cross (above) and horizontal (below) sections of light organ of *Luciola cruciata*. This is a variant of the Type 5 system of tracheal branching. After Okada (1935 b).

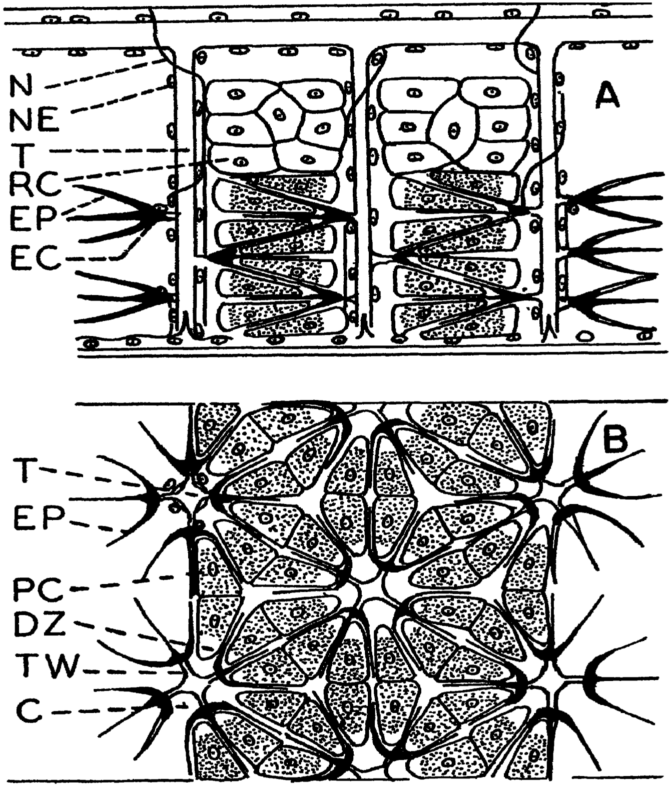


FIGURE 10. Diagram of cross (above) and horizontal (below) sections of the light organ of *Photinus* sp. (Presumably Korean). This is a variant of the Type 6 system of tracheal branching. After Okada (1935 b).

FIGURES 11-14 (*see opposite page*).

FIGURE 11. Composite diagram of end-cell and cytoplasmic structure from Figures 11, 16 and 20 of **Dahlgren** (1917). Lower end-cell more heavily impregnated with osmic acid than upper, and showing "fibers" in its cytoplasm and a "rounded body" (s) with postulated constrictor powers. Rounded photogenic granules in upper cells of male type; rod-shaped granules in lower photogenic cell of female type.

FIGURE 12. Cross section of light organ of *Pyrophorus*. After **Dubois**.

FIGURE 13. Diagram of horizontal section through light organ of *Photinus marginellus* to show "rosette" pattern of tracheolar .anastomoses between cylinders. Diameters of tracheoles exaggerated. Upper two cylinders show terminal brushes of tracheal twigs, lower shows nuclei of tracheal epithelium and end-cells; approx. x 1000. Partly after two figures of **Townsend** (1904).

FIGURE 14. Cross section of light organ of female of *Photuris pennsylvanica* from osmic acid preparation. After **Lund** (1911).

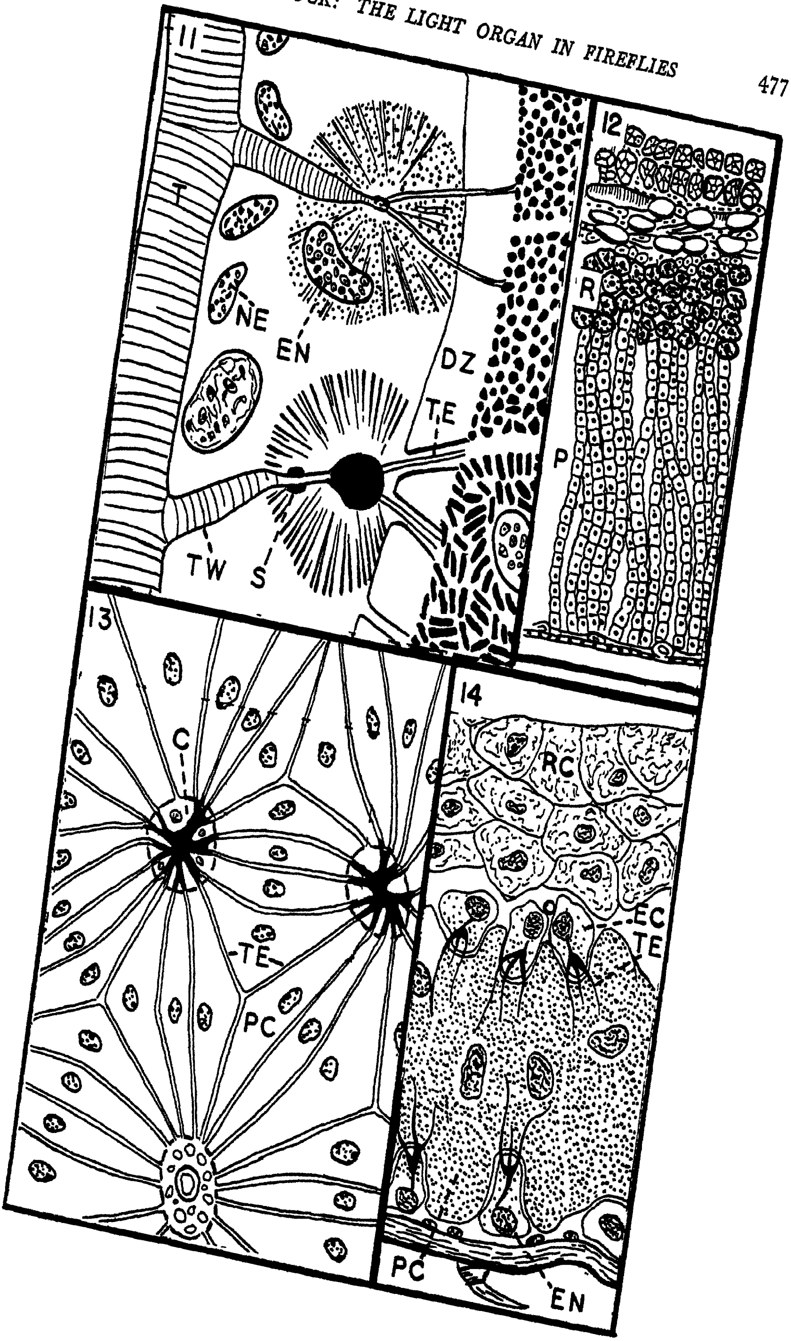

11
T
NE
EN
DZ
TE
TW
S
12
R
P
13
C
TE
PC
14
RC
EC
TE
PC
EN

PLATE 1

FIGURE 15 *Phengodes* sp Longitudinal section of lateral luminous organ, X180 Large oenocyte-like photogenic cells loosely aggregated in a fold of the external cuticle The photogenic cells are apparently slightly shrunken

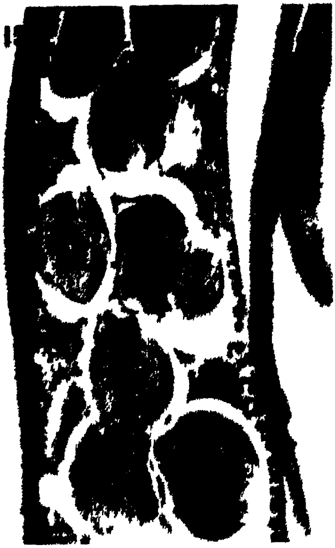
FIGURE 16 *Phrixothrix* sp Cross section of lateral luminous organ, X200 Organ is situated amid fat-bodies beneath the dorsolateral cuticle Clear spaces between photogenic cells are small tracheal branches

FIGURE 17 *Photuris pennsylvanica* Cross section of luminous organ of larva, X180 The two white dots at the apex of the photogenic layer are cross sections of tracheae Ventral cuticle slightly retouched

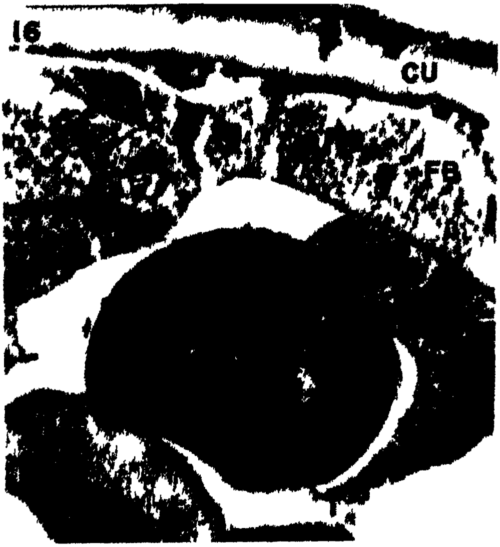
FIGURE 18 *Pyrophorus plagiophthalmus* Cross section of abdominal organ, X250

FIGURE 19 *Diphotus montanus,* female Cross section of part of luminous organ, X250 The separation between photogenic and reflector layers is an artifact

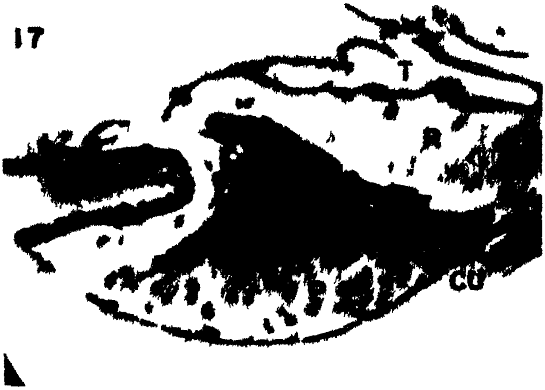
15



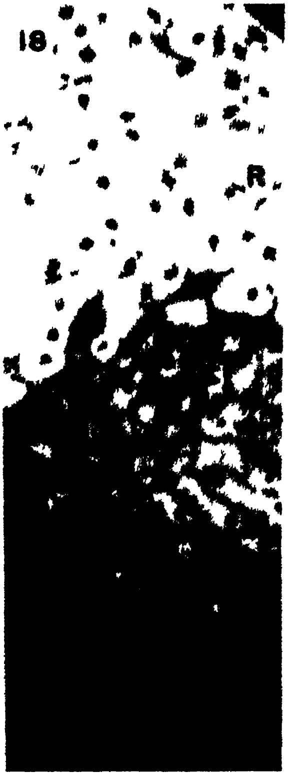
16



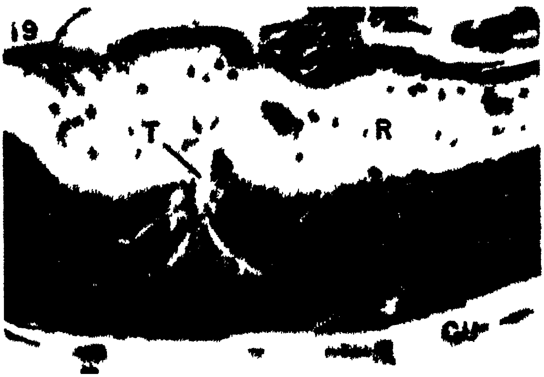
17



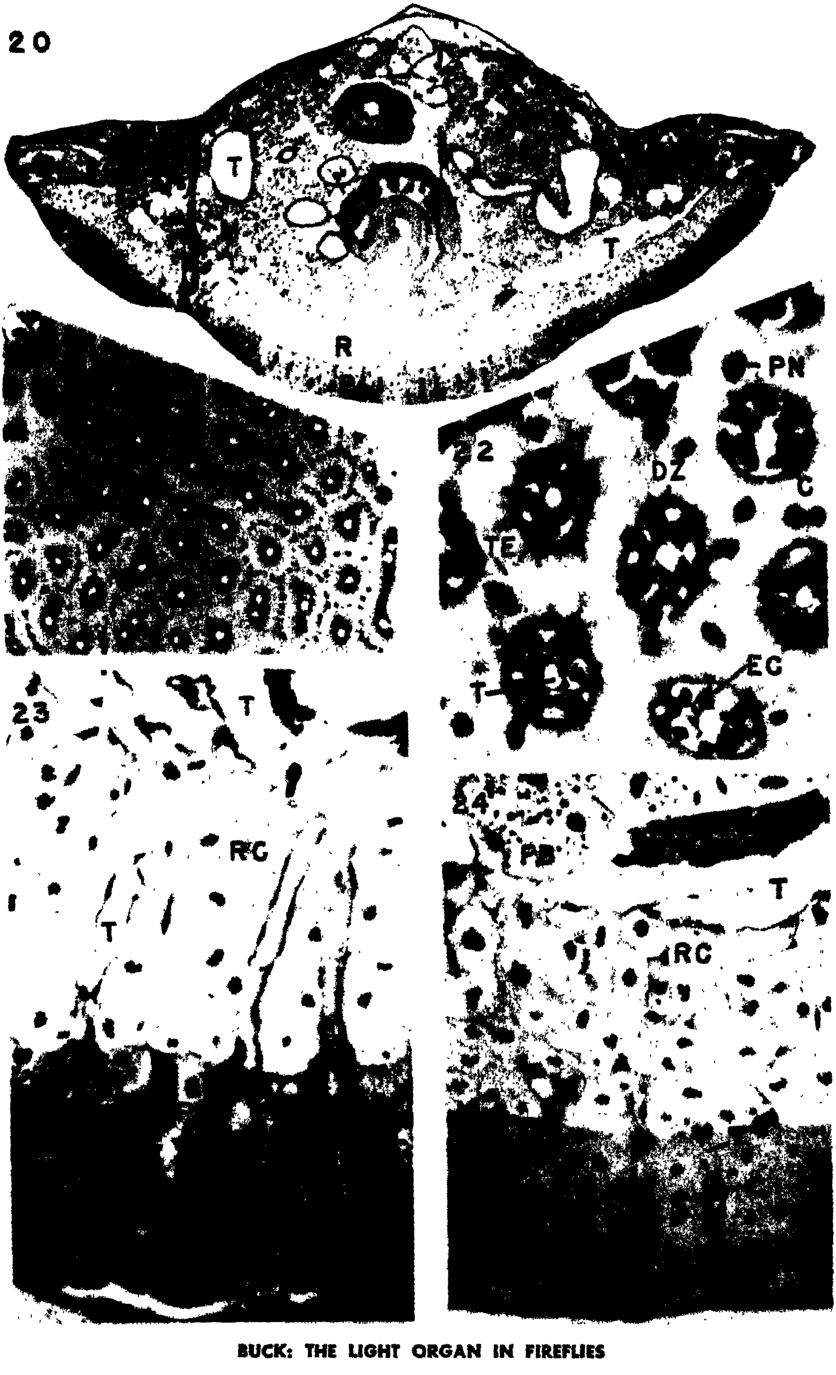
18



19



BUCK: THE LIGHT ORGAN IN FIREFLIES



BUCK: THE LIGHT ORGAN IN FIREFLIES

PLATE 2

FIGURE 20. *Photuris jamaicensis,* female. Cross section through entire abdomen, to show location of light organ on the ventral surface; X25. Note the high concentration of large tracheae on the internal surface of the reflector layer.

FIGURE 21. *Photinus xanthophotis catherinae*, male. Horizontal section through photogenic organ, to show the tracheal trunks arranged in triangular symmetry and the "rosettes" of photogenic cells around each; X96. The nuclei of the photogenic cells are arranged approximately equidistant from adjacent cylinders. The nuclei grouped close around each trachea belong mainly to tracheal end-cells, which are particularly abundant at the ventral surface of the organ (see also FIGURE 39). Compare with FIGURES 22, 29, and 36.

FIGURE 22. *Photinus euphotus*, male. Horizontal section through photogenic organ, showing "rosette" arrangement of photogenic cells around the cylinders. X540. This view also shows end-cells, tracheoles, cylinders, and the differentiated zone of the photogenic cytoplasm. Compare with FIGURES 21, 29, and 36.

FIGURE 23. *Photinus pyralis,* male. Cross-section through light organ: X250 This organ is characterized by large tracheal trunks, long narrow cylinders, photogenic and reflector layers of approximately the same thickness and number of cell layers (7 or more). One or two of the most internal of the photogenic cells show a lighter cytoplasm, such as has been described for "transition cells" between the photogenic and reflector layers (page 406). Compare with FIGURES 24, 25, 26, 31, 32, and 35.

FIGURE 24. *Photinus synchronans*, male. Cross-section through light organ: X250. This organ has narrow tracheae, and photogenic and reflector layers each about 5 cells thick. Compare with FIGURES 23, 25, 26, 31, 32, and 35.

PLATE 3

FIGURE 25 *Pl o u s a a cets* male Cross section through luminous organ ×250 In this species, the photogenic layer is usually thinner than the reflector, and the cylinders are often greatly flared at the ventral (and often the dorsal) surface of the photogenic layer Compare with FIGURES 23, 24, 26 31, 35, and especially 32

FIGURE 26 *Photinus pallens* male Cross section through photogenic organ ×230 Specially stained first method of Severinghaus to accentuate the differentiated zone of the photogenic cells Also shows cylinders expanded in the interior of the photogenic layer, and tracheoles Compare with FIGURES 2 , 24, 25, 31, 32, and 5

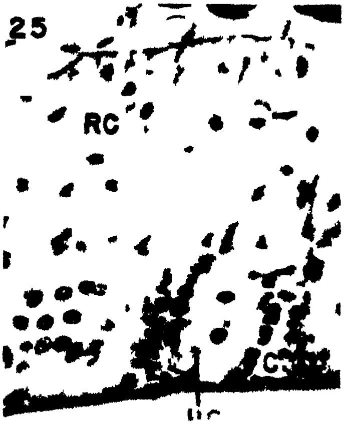
FIGURE 27 *Photinus pallens* male Three separated cylinders from osmic acid maceration preparation ×500 The end cells show varying degrees of impregnation from light (in the upper part of the right hand cylinder) to heavy (in the upper part of the left-hand cylinder) A few tracheoles shown in upper left and lower right Note the profusion of tracheal twigs and end cells

FIGURE 28 *Photinus pyralis* male Ventral surface view of photogenic organ which has been freshly peeled off ×500 The brushes of tracheae and twigs are especially profuse at the surface, since the cylinders usually flare there Dark 'wall" of cylinders may be the differentiated zone of the photogenic cytoplasm Compare with FIGURES 29 and 8

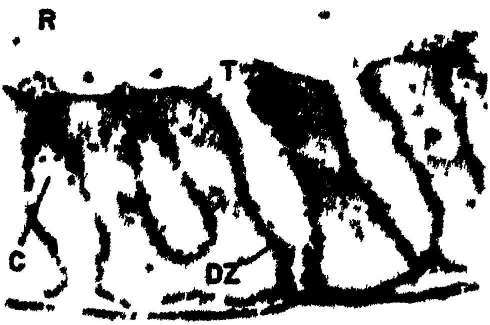
FIGURE 29 *Pl o ıs ı c l s* male Ventral view of photogenic organ which has been peeled off and partly dried ×790 Deeper focal level than FIGURE 28 Note tracheoles running between contiguous cylinders (several anastomoses are visible) Origin of tracheoles within cylinder is visible in the central cylinder and the one to the left of it Compare with FIGURES 28 and 38

FIGURE 30 *Photinus pallens* male Lateral view of terminal part of tracheal trunk of cylinder, from fresh smear ×500

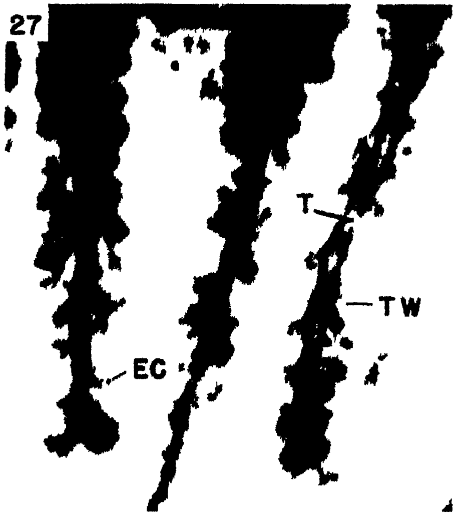
25



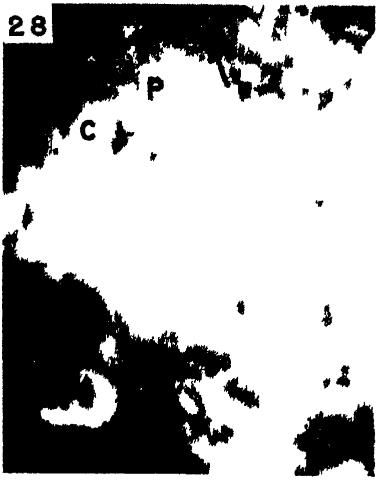
26

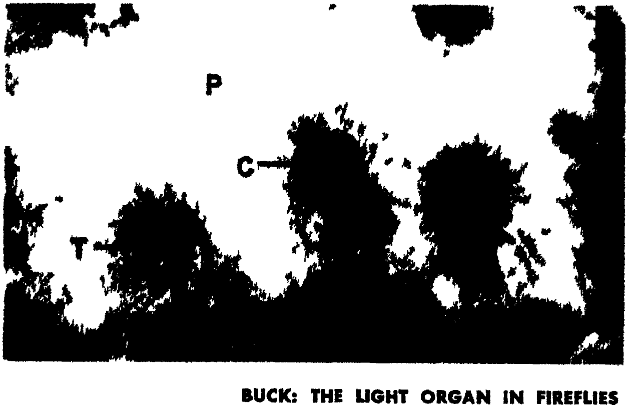


27



28





30

BUCK: THE LIGHT ORGAN IN FIREFLIES

31

32

33

34

PLATE 4

FIGURE 31 *Photinus evanescens montego* male Cross section through photogenic organ, X920 This species has a very thin photogenic layer, with extraordinarily broad cylinders At least 7 end cells show on the right edge of the cylinder in the center, and only three photogenic cells (judged from the number of nuclei) Note also the sharp delimitation of end cell cytoplasm and photogenic cytoplasm The differentiated zone of the photogenic cytoplasm, tracheal twigs, and several tracheoles show as well Nuclei of both end cells and tracheal epithelial cells occur in the cylinder Compare with FIGURES 23, 24, 25, 32, and 35

FIGURE 32 *Photuris pennsylvanica* male Cross section of photogenic organ X620 This figure shows the end cells with their tapering processes, but tracheoles are not visible In the lower right, where the end cells are cut transversely, note that each has 4 processes Note also end cell processes entering the photogenic cytoplasm from both surfaces of the photogenic layer Nuclei of end cells not distinguishable tracheal epithelial cells Compare with FIGURES 23, 24, 25, 31, and 35

FIGURES 33 and 34 *Photuris jamaicensis* male Horizontal section of the photogenic organ impregnated with silver nitrate X2300 These figures illustrate the delicate network which runs on the interfaces between the photogenic cells and binds contiguous tracheoles together

PLATE 5

FIGURE 35. *Photuris pennsylvanica,* female. Cross-section through photogenic organ; X830. This figure shows cell membranes (running vertically), along which appear cross-sections of tracheoles. Note also that the photogenic layer is only one cell thick and that its cells are very thin. No cylinder shows in this view. Compare with FIGURES 23, 24, 25, 31, and 32.

FIGURE 36. *Photinus pallens,* male. Ventral surface view of silver nitrate smear; X630. The membranes of the photogenic cells are seen running radially from the cylinders in a rosette pattern. Tracheoles viewed end-wise appear as small circles strung along the membranes. Compare with FIGURES 21, 22 and 29

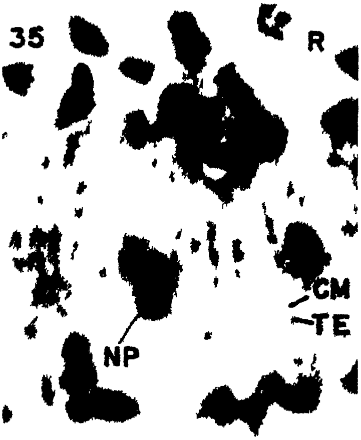
FIGURE 37. *Photinus pallens,* male. Osmic acid smear, showing isolated cylinder tissue with tracheae and lightly impregnated end-cells and tracheoles, X500 Origin of two tracheoles from one end-cell, and probable anastomoses, are visible.

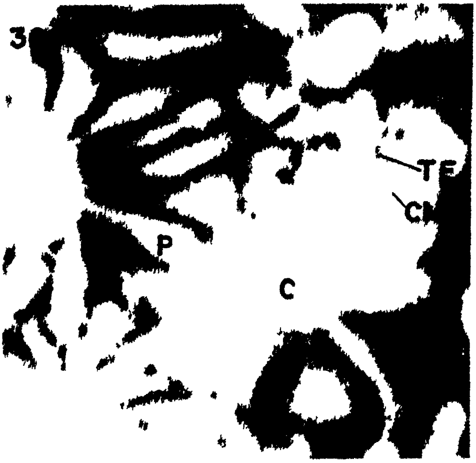
FIGURE 38. *Photinus pyralis,* male. Ventral view of partly dried photogenic organ, at level closer to surface than that shown in FIGURE 29; X790. This view shows the mode of origin of the tracheoles from the ends of the tracheal twigs within the cylinders. End-cells and other cytological details are invisible, since the preparation is unstained.

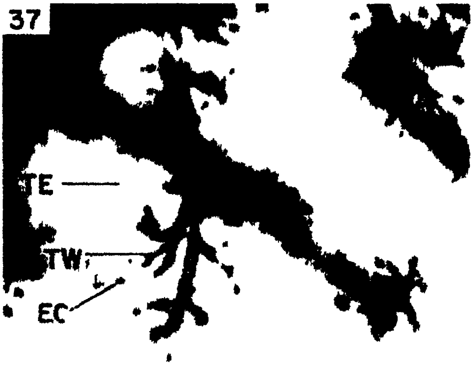
FIGURE 39. *Photinus xanthophotis catherinae,* male. Ventral surface view of organ peeled off after impregnation with osmic acid vapor: X425. This view shows the profuse brushes of end-cells which crowd the cylinders at the ventral surface where they flare. Note profuse tracheolar supply with many inter-cylinder anastomoses.

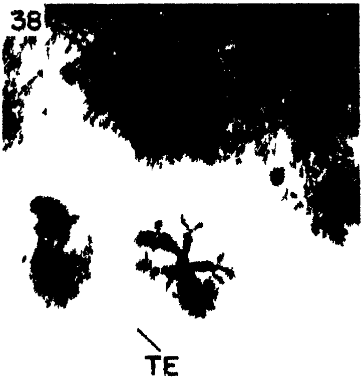
FIGURE 40. *Photuris jamaicensis,* male. Cluster of end-cells from terminal brush in flared ventral region of cylinder, from osmic acid smear; X500

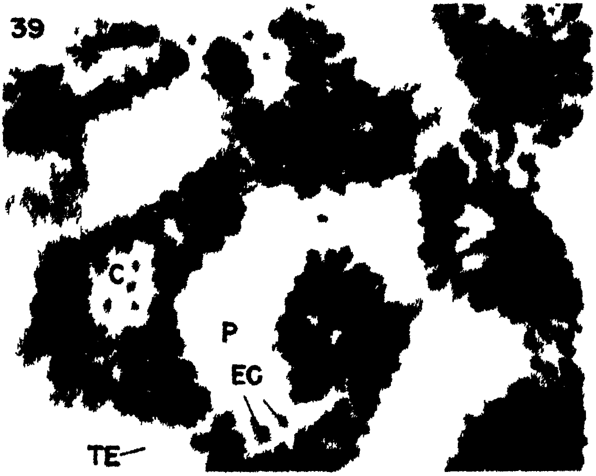
FIGURE 41 String galvanometer record of single spontaneous flash of *Photuris pennsylvanica.* Vertical axis, light intensity in arbitrary units, horizontal axis. time in 0.04 sec. intervals From Snell, 1932.

35


3


37


38


39


40

BUCK: THE LIGHT ORGAN IN FIREFLIES

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
VOLUME XLIX, ART. 4. PAGES 483-660
MAY 11, 1948

HEMORRHAGE*

CONTENTS

PAGE
Introductory Remarks. *By* GREGORY SHWARTZMAN 485
A Review of Some Basic Facts of Blood Coagulation. *By* JOHN H. FERGUSON 486
Heparin in Blood Clotting and Thrombosis. *By* C. H. BEST AND L. B. JAQUES 501
The Relationship of Vitamin K to Hemorrhage and Coagulation. *By* S. A. THAYER 518
Hemorrhagic Manifestations Observed in Experimental Deficiency of Pantothenic Acid, Choline, and Cystine. *By* PAUL GYŐRGY 525
The Effects of Hemorrhage on the Circulation. *By* DICKINSON W. RICHARDS, JR. 534
Experimental Studies on Traumatic and Hemorrhagic Shock. *By* MAGNUS I. GREGERSEN 542
Vasomotion in the Hemodynamics of the Blood Capillary Circulation. *By* ROBERT CHAMBERS 549
Reactions of Peripheral Blood Vessels in Experimental Hemorrhage. *By* B. W. ZWEIFACH, ROBERT CHAMBERS, R. E. LEE, AND C. HYMAN 553
Hepato-Renal Factors in Circulatory Homeostasis: III. The Influence of Humoral Factors of Hepato-Renal Origin on the Vascular Reactions to Hemorrhage. *By* EPHRAIM SHORR, B. W. ZWEIFACH, AND ROBERT F. FURCHGOTT 571
Effects of Hemorrhage on the Kidney. *By* DONALD D. VAN SLYKE....... 593
Metabolic Changes Associated with Hemorrhage. *By* ALFRED E. WILHELMI AND C. N. H. LONG.. 605
Certain Anatomo-Pathologic Aspects of Hemorrhage. *By* PAUL KLEMPERER 622
Hemorrhagic Manifestations of Bacterial and Virus Infections: Experimental Studies and Pathological Interpretations. *By* GREGORY SHWARTZMAN AND I. E. GERBER.............................. 627
Abnormal Hemorrhage with Normal Platelet Count and Normal Clotting. *By* RUSSELL L. HADEN, R. H. SCHNEIDER, AND LEE C. UNDERWOOD.. 641
Clinical Aspects of Hypoprothrombinemia. *By* CHARLES S. DAVIDSON AND HENRY J. TAGNON.. 647

---

* This series of papers is the result of a Conference on Hemoirhage, held by the Section of Biology of The New York Academy of Sciences on October 31, and November 1 and 2, 1946. Publication made possible by a contribution from the General Funds of the Academy.

COPYRIGHT 1948

BY

THE NEW YORK ACADEMY OF SCIENCES

INTRODUCTORY REMARKS

By Gregory Shwartzman

Division of Bacteriology, Laboratories of the Mount Sinai Hospital, New York, N. Y.

In opening the Conference on Hemorrhage, I should like to extend greetings to the members and guests of the Academy, and to welcome the distinguished scientists whose willingness to make contributions assures the success of our efforts. I should also like to express our gratitude to the Council of the New York Academy of Sciences, to Mrs. E. T. Miner, Executive Director, and Dr. Ross F. Nigrelli, Chairman of the Section of Biology, for making this conference possible; and to Drs. Charles H. Best, George R. Minot, and Donald T. Van Slyke for accepting the chairmanship of the sessions and advising on the organization of the program.

It becomes obvious upon looking over the long list of highly valuable conferences held by the Academy in the past that its basic policy is to aid in the correlation of scientific knowledge and to promote the integration of recent information into broader perspectives. The Conference on Hemorrhage which is offered at this time was planned with the same objective in mind. Due to the vastness of the field, the Conference could not be made as complete as desirable. Certain highly competent investigators who rightly belong here fail to appear on the program. Due to unforeseen circumstances, such pertinent investigations as those concerned with dicumarol and vitamin C are not covered. It is also realized that a conference of this sort remains incomplete without consideration of problems of erythroblastosis, cerebral hemorrhage, and the regeneration of formed elements of the blood and plasma following bleeding. The omissions are due, in part, to the fact that certain aspects have been already discussed at preceding conferences on Blood Grouping, Experimental Hypertension, and Some Aspects of Red Cell Production and Destruction, respectively.* Another factor is the fear that an attempt to present these topics comprehensively would have taken us into fields removed from the immediate scope of the Conference.

In keeping with the policy of the Academy, the number of papers has been limited in order to permit ample time for comments. It is hoped that intimate and vivid discussions will accompany the following presentations and permit the free exchange of views which has become a valuable tradition of these meetings.

---

* See the following publications of the Academy: Annals 46(9); Special Publications III; and Annals 48(7).

A REVIEW OF SOME BASIC FACTS OF BLOOD COAGULATION

By John H. Ferguson
Department of Physiology, School of Medicine,
University of North Carolina, Chapel Hill, North Carolina

The Fundamental Clotting Process. Human blood plasma contains about 0.28 gm. per cent of the protein fibrinogen.[16] The fundamental reaction of blood clotting is the conversion of fibrinogen from the state of colloidal solution ("sol") to a semi-solid fibrillar "gel", *fibrin.* Normally, this process requires thrombin, a natural specific coagulant which comes from another fraction of the plasma proteins. It is important, though not always easy, to distinguish between true fibrin and nonspecific denaturation of fibrinogen. Using the dark-field microscope to detect the characteristic fibrin "threads" (or "needles"), we find that papain enzyme gives clots indistinguishable from thrombin-fibrin, whereas the chemical ninhydrin produces only the nondescript appearances associated with protein denaturation.[36] Papain clots subsequently digest away (fibrinolysis). Thrombin-fibrin, however, is stable and does not usually show clot-retraction (syneresis), unless the reagents are contaminated with proteolytic enzymes of plasma or tissue origin.

Critique of Data on Mechanism of Fibrin Formation. *Papain* is a crystalline proteolytic enzyme from the paw-paw fruit, used commercially in "tenderizing" meat. Certain venoms[14] and some bacteria[87] are also said to clot fibrinogen directly. Ninhydrin is 1,2,3-indantrione (triketohydrindene) hydrate, especially associated with a color test for proteolysis developed by Abderhalden.[1] An allegedly similar pseudoclotting of fibrinogen is produced by chloramine-T, alloxan, and salicylaldehyde[7] (cf.[38]). The unsatisfactory nature of these and other studies in which protein denaturation occurs, makes it impossible to give much credence to the views (1) that fibrin formation is a "denaturation" process,[59] or (2) that certain oxidative reactions are fundamentally involved therein.[38] Recent work[34] has demolished the experimental basis for attributing some proteolytic action to thrombin. Analogies with papain, and possibly with rennin (although the clotting of milk casein is wholly unrelated to blood clotting), have no factual significance. The enzyme concept of thrombin,[15] therefore, is very insecure, and most of the facts fit equally well with a simpler approach *via* colloid chemistry.[32] A very recent development along these lines is supplied by Mommaerts.[48] He seeks to explain the mechanism of the thrombin-fibrinogen interaction on the basis of Bungenberg de Jong's *"coacervation"* concept.[5] A crude analogy may be made with reeds floating in a pond. Alone, or in isolated and haphazardly arranged bundles, these float in the suspending water. But if they are interlaced by cross-unions, a sufficiently firm mass is formed to hold the bulk of the water within its mesh. Now, there is good physical evidence that fibrinogen and fibrin are "fibrous" proteins composed of molecules which, though large in size, are extremely elongated and slender (filamentous).[12] Even in fibrinogen, these are aggregated into colloidal micelles, while in fibrin the oriented aggregates reach microscopically visible size in the form of the typical "needles" and "threads". If by some adsorptive or similar action, thrombin is able to affect the micelle orientation of fibrinogen and induce *"coacervation"*, the mesh structure and "gel" character of the resultant fibrin are readily explained. The latest values for the "isoelectric point" of fibrinogen,[29] at pH = 5.4, and for thrombin,[52] at pH = 4.4, are compatible with a role of electrochemical forces in the postulated adsorptive phenomena,[54] but many other considerations (*e.g.*, salts) and factors (*e.g.*, specific chemical bonds?) will need to be explored when our protein materials are sufficiently purified to justify the efforts that this will entail.

Experimental Method of Studying Clotting Reactions *in vitro*. Despite much progress in recent years,[26, 33, 34] few of the chemical agents concerned with the clotting process can as yet be isolated and determined with quantitative precision. We have long used a technique to circumvent this in the study of basic clotting reactions *in vitro*.[21] Firstly, experimental conditions must be strictly standardized, since these colloidal reactions are greatly influenced by temperature, pH, salt concentration, general dilution and particular amounts of specific agents, adsorption, and other surface phenomena. Examples of the last include (1) the use of gum acacia (see "fibrinoplastic" agents) in the Iowa technique for prothrombin assay;[55] and (2) the ability of "non-wettable" surfaces (*e g.*, silicone[43]) to retard coagulation of the blood. Secondly, each reagent and combination of reagents must be subjected to sensitive control tests, in order to rule out traces of impurities which may be significant in relation to their clotting properties. Finally, by confining experimental analysis to test-series conducted with the same batch of reagent, data are obtained which are of quantitative as well as of qualitative significance. The accurate measurement of time factors, especially clotting time (C.T.) is the key to these data, which will be illustrated with several examples.

RELATIONSHIP OF CLOTTING TIME AND RELATIVE THROMBIN CONCENTRATION (FIGURE 1). When a series of thrombin dilutions is tested on a given fibrinogen solution, it is found that the stronger the thrombin, the shorter is the clotting time. Under favorable and perhaps limited experimental conditions, we may even demonstrate a linear relation between the graphic plot of reciprocal clotting time against relative thrombin concentration. This is known as "the inverse law", but it is not necessary to go into its technicalities[39] to grasp the significance of the fundamental fact that a shorter clotting time means a more active agent, *i.e.*, more thrombin. Taking the undiluted thrombin as a standard (100), valid for the particular experiment, and referring to the C.T. values along the graph, the corresponding thrombic activity can be read off in relative percentages.[18]

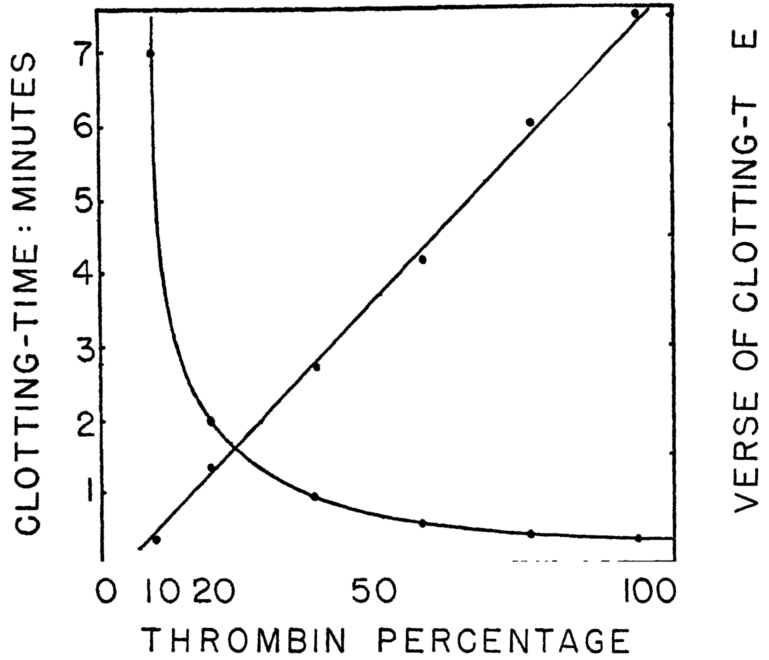


FIGURE 1. Relation of clotting-time (for pure fibrinogen) to varying concentration of thrombin.

The Formation of Thrombin from Prothrombin. It is only necessary to inject a potent thrombin solution, say into the ear vein of a rabbit, and to observe the prompt intravascular coagulation, in order to confirm the rather obvious fact that active thrombin is not present in normal circulating blood. What the plasma does contain is a protein precursor, prothrombin, which can be isolated and made to yield thrombin by appropriate activation procedures.

PROTHROMBIN ACTIVATION CURVES (FIGURE 2). In order to follow the activation process quantitatively, a mixture of prothrombin solution and activators is sampled at successive time intervals (incubation times). Each measured sample is added to a test fibrinogen solution and the clotting time noted. The graphical plot of these data is called a "prothrombin activation curve".[21] The shortening clotting times indicate more and more thrombin being formed. The slope and position of the curve express the rates of thrombin formation. The point at which the curve levels off, *i.e.*, shortest C.T., denotes complete (100 per cent) activation, a measure, therefore, of the relative amount of thrombin produced.

Inability to reproduce these curves exactly from one batch of materials to the next indicates that our reagents fall short of the requirements of pure biochemical entities. Nevertheless, the nature of the variability of the results offers a clue as to the type of impurities present. Thus, traces of natural activators, inhibitors, and proteolytic enzymes are suggested as common contaminants of prothrombin preparations. Technical difficulties in the purification of prothrombin are especially great, partly because of the problems associated with these impurities, and partly because of the very small amount (about 20 mg. per cent, according to the latest estimates[53]) of prothrombin present in plasma.

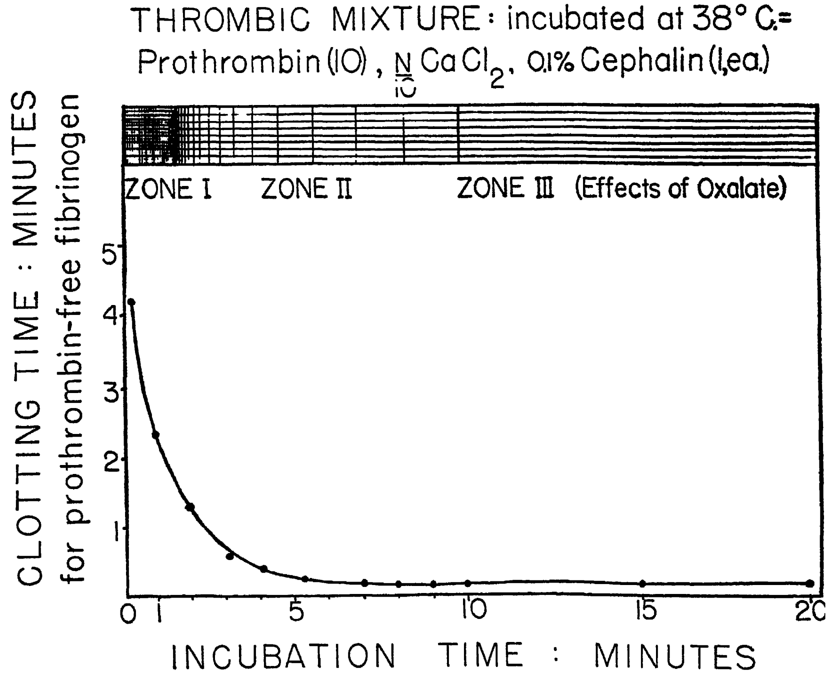


FIGURE 2. Prothrombin activation curve.

Activators of Prothrombin. The conversion of prothrombin to thrombin depends upon certain specific "activators":

CALCIUM. The blood calcium constitutes about 10 mg. per 100 cc. of serum. It is partly ionized (Ca^{++}) and partly in protein-bound and other nonavailable forms (in citrated plasma, for instance).[19] The exact nature of these complex equilibria, and often the time factors involved, have a definite bearing on certain coagulation problems, such as the need for a several-fold excess of the calcium ion depressing anticoagulants (oxalates, citrates, etc.) to prevent clotting, and the ability to restore clotting with much less than the chemical equivalent of calcium salt subsequently added.[49] The ability of trypsin to cause clotting in oxalated plasma is lost if the anticoagulant is increased.[85] Active Ca ions[46] are necessary for thrombin formation (first phase of clotting), but calcium is not essential for the thrombin-fibrinogen interaction (second phase of clotting). If too little calcium is present during the experimental activation of a prothrombin solution, the thrombin formation is slow and incomplete. Too much calcium is inhibitory, due to certain non-specific salt effects. The minimum Ca^{++} requirement for clotting is well below any blood level encountered clinically.[51] No clinical anomalies of blood coagulation can be attributed to variations in the blood calcium, and therapeutic claims for administering calcium in hemorrhagic disorders cannot be rationalized in terms of any known influence of calcium on the clotting mechanism. Calcium ion depressants, like oxalates and citrates, have long been used as anticoagulants. Added early enough (zone I, see FIGURE 2, in our studies[24] on prothrombin activation), they tie up the free Ca ions and prevent thrombin formation. In the final phase (zone III), after thrombin formation is complete, oxalates or citrates, apart from minor non-specific salt effects, have no action on thrombic activity. Even the complete removal of calcium by electrodialysis does not inactivate the thrombin. In an intermediate stage (zone II), however, treatment with a large excess of oxalate or citrate progressively inactivates the thrombic mixture. Re-addition of calcium salt restores the activation.

Our older experiments[20] on this point have recently been confirmed, with some minor differences, in some experiments on highly purified prothrombin (FIGURE 3). The facts are explained by the postulate of a calcium-containing intermediary complex (calcium-prothrombin-cephalin) appearing during thrombin formation.

THROMBOPLASTIC FACTORS. The best prothrombin preparations are not activated by calcium salt alone, or, at most, only very slowly and poorly. There is evidently need for an additional activator which we may describe as "thromboplastic" without commitment as to its nature or mode of action.[33]

Phospholipid. The phospholipid cephalin is one, but not necessarily the only, such factor. Cephalin is abundant in plasma, including that of hemophiliacs,[17] and can be isolated from the brain and other tissues in a state of reasonable chemical purity. We have always found that crude, and (in recent tests) highly purified, prothrombin is completely converted into thrombin within a few minutes after the addition of optimal amounts of calcium salt and pure cephalin. Inadequate amounts of cephalin result in slow and incomplete thrombin formation. The final thrombin may be prepared completely free from phospholipid.[2] Owing to the avidity with which the ubiquitous phospholipids and metal cations (*e.g.*, Ca), combine with proteins, no one has yet reported the preparation of prothrombin free from these activator agents. Nevertheless, there is good evidence that the protein-bound forms of these agents do not directly participate, and that to form thrombin they must be present in the "free" or active state. The data, therefore, indicate that the phospholipid (cephalin), like calcium, participates *via* an intermediate, *i.e.*, prothrombin-calcium-cephalin complex or compound.

Thromboplastin (thrombokinase). While non-thromboplastic phospholipid-protein compounds undoubtedly exist, recent workers[8] have

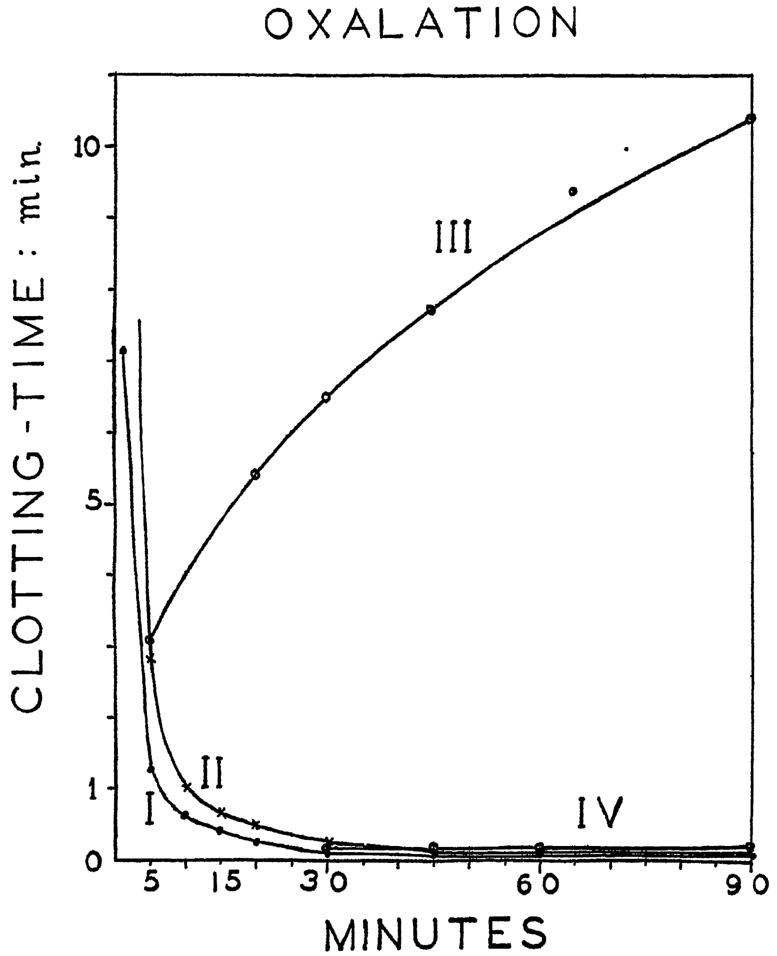


FIGURE 3. Effects of oxalation on thrombic activity. (Experiments with purified prothrombin at 25° C.)
I. prothrombin activation without oxalate (control).
II. prothrombin activation, with oxalate added to fibrinogen (control of second phase).
III. thrombic mixture oxalated 5 minutes from start (zone II).
IV. thrombic mixture oxalated 30 minutes from start (zone III).

isolated "thromboplastic lipoprotein", of which the essential phospholipid component may not consist entirely of cephalin.[6] Our work, to date, has included only the crude aqueous tissue extracts long known as thromboplastin or thrombokinase, and study of these has led us in another direction.[35] Briefly summarizing, these agents require Ca^{++}, are more active than cephalin or their isolated total P-lipids, and, unlike cephalin, they can (with Ca^{++}) form thrombin in the presence of heparin (see below) or in hemophilic plasma.

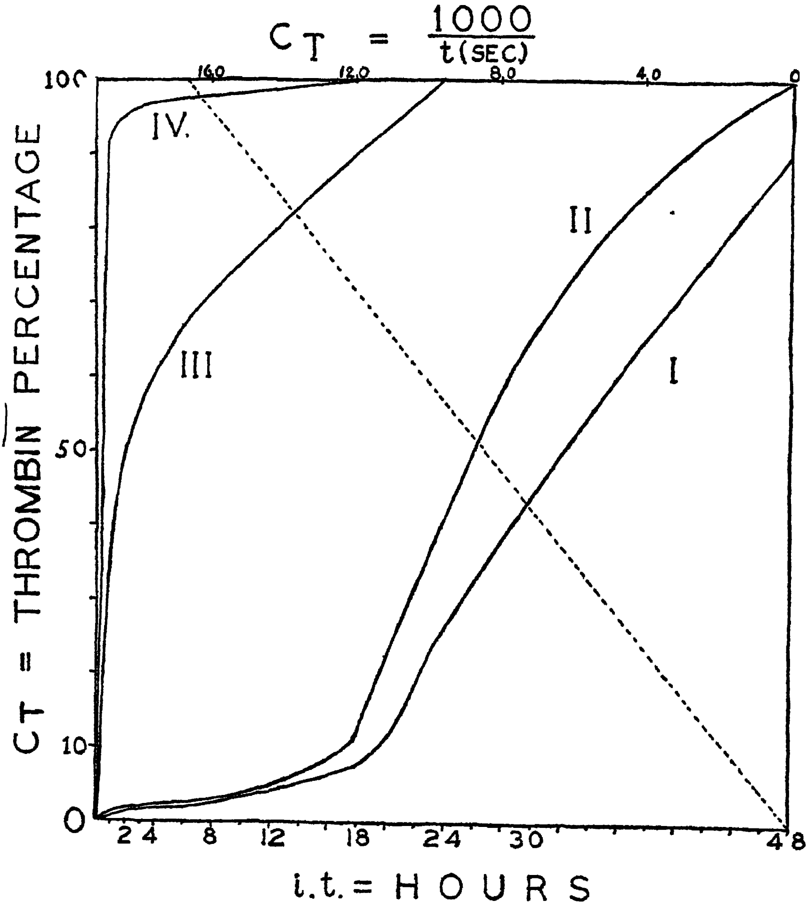


FIGURE 4. Effects of trypsin on thrombin formation. (Experiments with purified prothrombin and fibrinogen.)

Prothrombin activation curves (cf. FIGURE 2) for incubation times (i.t.) stated, at 25°C. pH = 7.7 (borate buffer). By reference to dotted line (inverse of clotting-times, *t*, of dilutions of fully activated mixture III) thrombic activity (C_T) is expressed in percentages.

Mixture I: prothrombin + calcium salt (alone).
Mixture II: prothrombin + Ca + trypsin.
Mixture III: prothrombin + Ca + brain thromboplastin.
Mixture IV: prothrombin + Ca + thromboplastin + trypsin.

Tryptases. These additional thromboplastic actions (of thromboplastin) can be reproduced *in vitro* by adding a small quantity of crystalline trypsin (pancreatic enzyme) to the simpler Ca + cephalin thromboplastic system.[35] Intravascular coagulation and anaphylactoid shock are dangers which preclude the use of trypsin *in vivo*.[34] Recent unpublished experiments (FIGURE 4) with purified prothrombin confirm an important conclusion previously reached, namely, that trypsin is not a thromboplastic agent in its own right, but merely catalyzes the actions of cephalin and calcium when thrombin is being formed in the presence of interfering proteins. Purified prothrombin is well activated by calcium and cephalin in the absence of proteolytic enzyme. But in the presence of other proteins that divert some of the cephalin and calcium, trypsin attacks the proteins and so mobilizes the activators for thrombin formation. Obviously, this is of great importance in the natural clotting system, and natural proteolytic enzymes, *e.g.*, serum-tryptase, are factors upon which to base an explanation of the initiation of clotting in shed blood.[31]

The coagulation delay in hemophilia can be counteracted *in vitro* by (1) tissue thromboplastin, but not significantly by pure cephalin,[23] (2) trypsin (thromboplastic enzyme),[23, 24] or (3) "antihemophilic globulin," in the form of Fraction I of the plasma proteins prepared by the Harvard laboratories.* The last is safe and effective *in vivo*, but not the first two. Whether or not these data will lead to an interpretation involving the (thromboplastic) protease system[18, 56] awaits further experimental elucidation.

Lytic Phenomena. Natural tryptases have ordinary proteolytic actions, including fibrinolysis and fibrinogenolysis (which we use for an enzyme assay method[30] sensitive to 1:1,000,000 of a trypsin standard), also prothrombinolysis and thrombinolysis. Although essentially independent, these do encroach upon the coagulation problem, *inter alia*, in (1) preparation of protein clotting factors, (2) clot retraction and fibrinolysis, (3) non-coagulability of menstrual blood, and (4) fibrin resolution within the body under pathological conditions.[34] †

Clot-Inhibitors. Ill-defined classical terms like "antithrombin",[3, 4] "metathrombin",[60] etc., should be dismissed in favor of a return to the experimental analysis of clot-inhibitory mechanisms.[39] At the present state of our knowledge, it is much safer to stick to non-committal descriptive adjectives like "antithrombic" and "antiprothrombic", signifying no more than the fact that the respective inhibition occurs in the second or in the first phase of the coagulation reactions. Such data as approach the problem along the lines of colloid chemistry, with particular reference to adsorption phenomena, offer definitely useful leads.

---

* Edsall, J. T. Adv. Protein Chem. 3: 383. 1947.

† With the aid of a grant from the John and Mary R. Markle Foundation, we are currently investigating the enzyme problem, and it may be mentioned that there are already experimental leads pointing to (1) an inactive enzyme precursor (tryptogen); (2) a kinase-type activator (tryptokinase), an example of which is streptokinase,[11] long miscalled streptococcal fibrinolysin;[47] (3) inhibitors (antitryptase), of which we have found the pancreatic trypsin-inhibitor (a crystalline polypeptide) inhibitory to blood coagulation.[28] The same is also true of a new crystalline trypsin-inhibitor from soy-beans.[45]

The best known natural clot-inhibitor is heparin,[27] believed to originate in the metachromatic-staining granules of the tissue basophils or Ehrlich "mast-cells".[58] The singular is used for convenience, although it really represents a class of substances, having the general composition of mucoitin-polysulfuric-esters (*i.e.*, complex carbohydrate derivatives), the molecular building stones of which are (1) glycuronic acid, (2) glucosamine, (3) acetic acid (?), and (4) sulfuric acid (ester-linked).[44]

Heparin inhibits clotting both *in vitro* and *in vivo*.[42] Its modes of action are complex and depend upon the cooperation of a co-factor (heparin-complement), which is not yet identified but seems to be a component of the crude "albumin" fraction of the plasma proteins.[50] Heparin, *plus* co-factor, is both antithrombic and antiprothrombic. Salmine (protamine) reverses these inhibitions both *in vitro* and *in vivo*.[9] By itself, salmine is antiprothrombic in the first phase and clot-aiding (fibrinoplastic) in the second.[25] The term "fibrinoplastic" merely denotes a non-specific facilitation of fibrin formation, shown by many colloids. In its first-phase (antiprothrombic) actions, heparin (*plus* co-factor) shows quantitative antagonisms toward the thromboplastic factors.[22] Part of this may be a deviation of phospholipid[10] probably to the heparin co-factor or other coagulation-inert proteins. Heparin alone, in unphysiologically large amounts, inhibits the ordinary proteolytic actions of trypsin,[40, 41] but it is not yet certain that this applies to inhibition of the "thromboplastic" action of natural (serum-) tryptase under physiological conditions. *In vivo,* heparin prevents the intravascular coagulation but not the anaphylactoid toxic effects of injected trypsin.[57] It is reasonable to conclude that a major part of the antiprothrombic action of heparin is directed against the activators of prothrombin. The possibility that some direct alteration of the prothrombin is also involved is suggested by some experiments, but requires confirmation on purer prothrombin.

Heparin also lessens the agglutination and breakdown of blood platelets, but many facts, including (1) their quantitative insignificance, as compared with plasma sources of the clotting factors, and (2) the absence of coagulation delay in thrombocytopenic purpuras, argue against any real importance of the platelets in blood coagulation. Clotting is essentially a plasma phenomenon, which contributes one of the several defense mechanisms which the body can call upon when faced with the problem of hemorrhage.[24]

SUMMARY

THE REACTIONS OF BLOOD CLOTTING

First Phase: Conversion of Prothrombin to Thrombin

PROTHROMBIN PHOSPHOLIPID CALCIUM IONS

PROTHROMBIN–CEPHALIN–CALCIUM
"Intermediary" Complex

THROMBIN

Note: (1) Fully formed thrombin can be obtained free from calcium and phospholipid.
(2) Small amounts of trypsin enzyme "catalyze" the above reaction. Natural serum-tryptase is probably essential for the reaction under ordinary blood clotting conditions.
(3) Purified prothrombin, treated to minimize the effects of contaminating traces of accessory factors, *also* needs "accelerator globulin" (see *Addendum)* for rapid completion of thrombin formation.

Second Phase: Conversion of Fibrinogen (Sol) to Fibrin (Gel)

FIBRINOGEN THROMBIN

FIBRIN

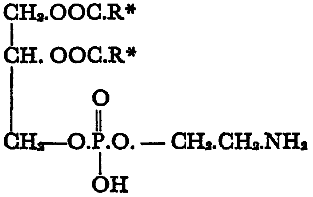
Note: (1) These collodial reactions are influenced by:
(a) temperature;
(b) pH;
(c) salt concentration;
(d) concentration (dilution) of specific agents;
(e) adsorption and other collodial factors, *e.g.*, "wetting".
(2) The blood plasma is a potential source of all the clotting factors, but platelets and damaged cells and tissues may help, especially with regard to "thromboplastic" and "fibrinoplastic" factors.

CALCIUM IN CLOTTING REACTIONS

(1) Ca ions ordinarily essential.
(2) Ca not needed for thrombin-fibrinogen reaction (second phase).
(3) Excess Ca salt delays clotting: both phases.
(4) Insufficient Ca^{++} slows and lessens thrombin production (first phase).
(5) *Ca ion depressants* (oxalates, citrates, etc.) in first phase:
Zone I: prevent thrombin formation,
Zone II: progressively inactive "intermediary",
Zone III: do not inactivate final thrombin.
(6) Decalcification and recalcification involve complex equilibration of free and bound Ca, needing time.
(7) No clinical variations of blood calcium significantly affect the clotting mechanism.
(8) Ca^{++}, alone, is insufficient for thrombin formation.

PHOSPHOLIPID IN THROMBIN FORMATION

(1) Natural cephalin is a "thromboplastic" agent, inactive alone, but active along with Ca^{++}.
(2) Prothrombin (crude), containing P-lipid but not (or poorly) activated by Ca alone, is quickly and completely converted to thrombin by Ca + cephalin.
(3) Varying the cephalin concentration determines (a) rate, and (b) amount of thrombin formation.
(4) Pure thrombin is phosphorus-free.
(5) Cephalin is normally bound to protein, but can be deviated (freed) by heparin.
(6) Cephalin is normal in amount and readily bound by protein in hemophilic plasma.
(7) The formula for cephalin is:



$$\begin{array}{l} \mathrm{CH_2.OOC.R^*} \\ | \\ \mathrm{CH.\ OOC.R^*} \\ | \\ \mathrm{CH_2{-}O.}\underset{\overset{|}{\mathrm{OH}}}{\overset{\overset{\mathrm{O}}{\|}}{\mathrm{P}}}\mathrm{.O.{-}CH_2.CH_2.NH_2} \end{array}$$

i.e., consisting of: 1. glycerol;
2. fatty acids (R*);
3. phosphoric acid;
4. the nitrogenous base, ethanolamine.

It is present in blood plasma and all cells and tissues, in fairly firm combination with proteins.

THROMBOPLASTIN (THROMBOKINASE) IN THROMBIN FORMATION

In general, thromboplastins (aqueous tissue extracts):

(1) need calcium ions to convert prothrombin to thrombin,
(2) are more potent than cephalin or their isolated phosphatides,
(3) act in the presence of considerable heparin,
(4) reduce the clotting time of hemophilic plasma (*in vitro*),
(5) often have demonstrable proteolytic effects.

Specifically, lung lipoprotein (Chargaff *et al.*):

(6) is a very active "thromboplastic" agent,
(7) on analysis, yields
    (a) cephalin and other lipids (? unknown thromboplastic phosphatide),
    (b) a de-fatted protein residue which is non-thromboplastic and may, *e.g.* with heparin, participate in clot inhibition.

TRYPTASES IN RELATION TO CLOTTING

(1) Tryptases (Oppenheimer: trypsin-like proteolytic enzymes) are ubiquitous, *e.g.*, serum-tryptase, cell-tryptases.
(2) Pancreatic trypsin (crystalline)
    (a) *in vivo*, produces intravascular clotting and shock;
    (b) *in vitro*, clots plasma (1) oxalated (or citrated), (2) heparinized, (3) hemophilic.
(3) Its mode of action is "thromboplastic" (Eagle, Ferguson).
(4) Current experimental data (Ferguson) suggest
    (a) details of trypsin action vary with purity of prothrombin;
    (b) trypsin aids but is not essential for thrombin-forming action of calcium cephalin (or thromboplastin);
    (c) trypsin is not thromboplastic in the absence of these factors.

TRYPSIN AND (POSTULATED) TRYPTASE SYSTEMS

Precursor (zymogen):	Pancreatic trypsinogen (crystalline protein)	Plasma-tryptogen
Activator (kinase):	(1) enterokinase (2) mold kinase	Tryptokinase (1) Streptokinase (strep. "fibrinolysin")
Active enzyme:	Trypsin (crystalline)	Serum-tryptase
Inhibitors (anti-tryptic):	Trypsin-inhibitors (1) pancreatic crystalline polypeptide (2) crystalline soybean inhibitor	Antitryptase

HEPARINS: MUCOITIN-POLYSULFURIC ESTERS

Anticoagulant *in vitro* and *in vivo*

Constitution: (1) glycuronic acid, (2) glucosamine, (3) acetic acid (?), (4) sulfuric acid (ester).

Modes of action: heparin + co-factor (heparin-complement).
In first phase —(a) anti-thromboplastic: (1) deviates cephalin,
(2) inhibits tryptase (?).
(b) anti-prothrombic (? lessens thrombin yield).
In second phase— anti-thrombic.
In whole blood — preserves and lessens agglutination of platelets.

Note: (a) acidic heparin combines with basic protein or protamine.
(b) protamine (*e.g.*, salmine) reverses heparin effects, *in vitro* and *in vivo*.
(c) salmine, alone, is inhibitory in first phase but aids second phase of clotting.

ADDENDUM

Since this review has gone to press, evidence from several laboratories advances the highly significant idea that a hitherto unrecognized clotting-factor ("factor 5" of Owren,[61] "labile factor" of Quick,[62] "accelerator globulin" of Ware, Seegers, *et al.*[63]) is an important *accessory factor* in the conversion of prothrombin to thrombin. Owren[61] describes a clinical case of severe hemorrhagic disorder apparently correlated with a specific deficiency of this new factor. When plasma or isolated prothrombin is stored under temperature and other conditions favorable to deterioration, loss of activity of the accessory factors as well as of the prothrombin itself, must receive due consideration.[64]

BIBLIOGRAPHY

1. **Abderhalden, E.**
1922. Die Abderhaldensche Reaktion. Fifth Edition. J. Springer. Berlin.
2. **Astrup, T., & S. Darling**
1941. Acta Physiol. Scand. **2**: 22.
3. **Astrup, T., & S. Darling**
1942. Acta Physiol. Scand. **4**: 293.
4. **Astrup, T., & S. Darling**
1943. Acta Physiol. Scand. **5**: 13.
5. **Bungenberg de Jong, H. G.**
1936. La Coacervation. J. Hermann. Paris.

6. **Chargaff, E.**
1944. J. Biol. Chem. **155**: 387.
7. **Chargaff, E., & A. Bendich**
1943. J. Biol. Chem. **149**: 93.
8. **Chargaff, E., D. H. Moore, & A. Bendich**
1942. J. Biol. Chem. **145**: 593.
9. **Chargaff, E., & K. B. Olson**
1937. J. Biol. Chem. **122**: 153.
10. **Chargaff, E., M. Ziff, & S. S. Cohen**
1940. J. Biol. Chem. **136**: 257.
11. **Christensen, L. R., & C. M. MacLeod**
1945. J. Gen. Physiol. **28**: 559.
12. **Cohn, E. J.**
1945. Sci. Progress **4**: 273.
13. **Eagle, H.**
1935. J. Gen. Physiol. **18**: 531
14. **Eagle, H.**
1937. J. Exp. Med. **65**: 613.
15. **Eagle, H.**
1941. Macleod's Physiology in Modern Medicine. Ninth Edition: 313. Edited by P. Bard, C. V. Mosby. St. Louis.
16. **Edsall, J. T., R. M. Ferry, & S. H. Armstrong, Jr.**
1944. J. Clin. Invest. **23**: 557.
17. **Erickson, B. N., & J. H. Ferguson**
1940. Proc. Soc. Exp. Biol. & Med. **45**: 579.
18. **Feissly, R.**
1941. Helv. Med. Acta **8**: 823.
19. **Ferguson, J. H.**
1936. Physiol. Rev. **16**: 640.
20. **Ferguson, J. H.**
1937. Am. J. Physiol. **119**: 755.
21. **Ferguson, J. H.**
1938. J. Lab. & Clin. Med. **24**: 273.
22. **Ferguson, J. H.**
1939. Proc. Soc. Exp. Biol. & Med. **42**: 33.
23. **Ferguson, J. H.**
1939. Am. J. Physiol. **126**: 669.
24. **Ferguson, J. H.**
1940. J. Lab. & Clin. Med. **26**: 52.
25. **Ferguson, J. H.**
1940. Am. J. Physiol. **130**: 759.
26. **Ferguson, J. H.**
1940. Ann. Rev. Physiol. **2**: 71.
27. **Ferguson, J. H.**
1941. J. Lab. & Clin. Med. **26**: 1559.
28. **Ferguson, J. H.**
1942. Proc. Soc. Exp. Biol. & Med. **51**: 373.
29. **Ferguson, J. H.**
1942. J. Gen. Physiol. **25**: 607.
30. **Ferguson, J. H.**
1943. Proc. Soc. Exp. Biol. & Med. **52**: 243.
31. **Ferguson, J. H.**
1943. Science **97**: 319.
32. **Ferguson, J. H.**
1944. Colloid Chemistry: Theoretical and Applied. **5**: 951. Edited by J. Alexander. Reinhold. New York.

33. **Ferguson, J. H.**
1946. Howell's Textbook of Physiology. Fifteenth Edition: 548, 586. Edited by J. F. Fulton. W. B. Saunders. Philadelphia.

34. **Ferguson, J. H.**
1946. Ann. Rev. Physiol. **8**: 231.

35. **Ferguson, J. H., & B. N. Erickson**
1939. Am. J. Physiol. **126**: 661.

36. **Ferguson, J. H., & P. H. Ralph**
1943. Am. J. Physiol. **138**: 648.

37. **Fredericq, P.**
1942. Arch. Internat. Physiol. **52**: 73.

38. **Gerber, C. F., & E. W. Blanchard**
1945. Am. J. Physiol. **144**: 447.

39. **Glazko, A. J., & J. H. Ferguson**
1940. J. Gen. Physiol. **24**: 169.

40. **Glazko, A. J., & J. H. Ferguson**
1940. Proc. Soc. Exp. Biol. & Med. **45**: 43.

41. **Horwitt, M. K.**
1944. J. Biol. Chem. **156**: 427.

42. **Jaques, L. B., A. F. Charles, & C. H. Best**
1938. Acta Med. Scand. Suppl. **90**: 190.

43. **Jaques, L. B., E. Fidlar, E. T. Feldsted, & A. G. MacDonald**
1946. Fed. Proc. No. 1(2). **5**: 52.

44. **Jorpes, J. E.**
1939. Heparin: Its Chemistry, Physiology, and Application in Medicine. Oxford Univ. Press. London.

45. **Kunitz, M.**
1946. J. Gen. Physiol. **29**: 149.

46. **Lebel, H., F. Schönheyder, & J. Muus**
1938. Skand. Arch. Physiol. **78**: 179.

47. **Loomis, E. C., & R. M. Smith**
1946. J. Biol. Chem. **163**: 767.

48. **Mommaerts, W. F. H. M.**
1946. J. Gen. Physiol. **29**: 103, 113.

49. **Nordbö, R.**
1936. Skand. Arch. Physiol. **75**, Suppl. 11.

50. **Quick, A. J.**
1944. Physiol. Rev. **24**: 297.

51. **Ransmeier, J. C., & F. C. McLean**
1938. Am. J. Physiol. **121**: 488.

52. **Seegers, W. H.**
1940. J. Biol. Chem. **136**: 103

53. **Seegers, W. H., E. C. Loomis, & J. M. Vandenbelt**
1945. Arch. Biochem. **6**: 85

54. **Seegers, W. H., M. Nieft, & E. C. Loomis**
1945. Science **101**: 520.

55. **Seegers, W. H., & H. P. Smith**
1942. Am. J. Physiol. **137**: 348.

56. **Tagnon, H. J., C. S. Davidson, & F. H. L. Taylor**
1943. J. Clin. Invest. **22**: 127.

57. **Wells, J. A., C. A. Dragstedt, J. A. Cooper, & H. C. Morris**
1945. Proc. Soc. Exp. Biol. & Med. **58**: 57.

58. **Wilander, O.**
1938. Skand. Arch. Physiol. **81**, Suppl. 15.

59. **Wöhlisch, E.**
1940. Ergebn. Physiol. **43**: 174.

60. **Wöhlisch, E.**
1943. Biochem. Z. **316:** 295.
61. **Owren, P. A.**
1947. The Coagulation of Blood. Investigations on a New Clotting Factor. J. Chr. Gundersen. Oslo.
62. **Quick, A. J.**
1947. Am. J. Physiol. **151:** 63.
63. **Ware, A. G., M. M. Guest, & W. H. Seegers.**
1947. J. Biol. Chem. **169:** 231.
64. **Munro, M. P., & F. L. Munro.**
1947. Am. J. Physiol. **150:** 409.

DISCUSSION OF THE PAPER*

Dr. A. L. Copley *(New York University, New York, N. Y.):*

I believe that we should differentiate between (a) fibrin formation, and (b) gelation which comprises the third phase of blood coagulation. Fibrin precipitation without gelation can be observed in samples of blood or plasma made less coagulable with sodium heparin, physiologic saline, and silicone-treated surfaces recently developed by Jaques *et al.*,[1] or in hemophilic blood and plasma. Gelation is not necessarily dependent upon fibrin formation, as is generally believed. In *Limulus* blood which does not contain fibrinogen, I observed two phases of clotting: (a) the agglutination of amoebocytes, and (b) the gelation of blood, thus corroborating earlier findings by Leo Loeb. To my knowledge, the blood of *Limulus polyphemus* represents the most rapid clotting system of native blood. We found agglutination times to vary from 6 to 20 seconds. The initial gelation was noticed as early as 24 seconds and varied up to 55 seconds.[2] Calcium may accelerate gelation, as found in certain model tests using prothrombin, thromboplastin, and fibrinogen.

References

1. **Jaques, L. B., E. Fidlar, E. T. Feldsted, & A. G. MacDonald**
1946. Canad. Med. Assoc. J. **55:** 26.
2. **Copley, A. L.**
1947. Fed. Proc. **6:** 90.

---

* See also discussion by Dr. Moolten on pages 512-515.

HEPARIN IN BLOOD CLOTTING AND THROMBOSIS

By C. H. Best and L. B. Jaques*

Department of Physiology, and Banting and Best Department of Medical Research, University of Toronto, Toronto, Canada

To continue the discussion of the biochemical aspects of hemorrhage, we have been asked to deal with the subject of heparin. In only one pathological condition, namely, severe anaphylactic shock, is heparin found in sufficient amounts in the blood to increase the danger of hemorrhage. Even when the heparin content of the blood has been raised by injection of the material, it is surprisingly easy to control hemorrhage in the normal individual. There remains, however, the problem of changes in the physiology of heparin during hemorrhage, which has not yet been properly studied. It is hoped that this discussion may throw some light on how to attack this problem.

Heparin and Clotting *in Vitro*. Considering the effect on clotting when added to blood *in vitro*, heparin is most remarkable in its anticoagulant activity. One part of heparin will prevent the clotting of 100,000 times its weight of blood. No other substance is as effective in this regard. Hence, starting with the original investigations of Professor Howell, many investigators have attempted to elucidate the mechanism of this action. As Brinkhous and collaborators[1] have emphasized, prothrombin is present unchanged in heparinized blood, and hence, heparin must act in the first stage of clotting. Heparin also has an action on the second stage of clotting, its so-called antithrombin action. Due to the difficulties involved in studies of the first stage, this antithrombic activity has been more thoroughly investigated. As will be indicated later, there is a possibility that this antithrombin action may be an entirely separate property, and that argument by analogy applied to the action on the first stage of clotting may, therefore, be quite misleading. However, as Howell originally showed, heparin acts as an antithrombin only upon addition of a co-factor present in plasma. The conclusion of Professor Howell that this co-factor is distinct from plasma antithrombin has recently been completely confirmed by Astrup and Darling,[2] and by Feisly and Enowicz.[3] Dr. Ferguson has already referred to the significance of plasma trypsin in blood coagulation. Horwitt[4] and Rocha e Silva and Andrade[5] have shown that heparin inhibits crystalline trypsin and also serum protease. However, its activity in this regard is too low to explain its anticoagulant action. Rocha e Silva and Andrade have shown that heparin also inhibits the activation of plasma trypsin by trichloracetic acid.

---

* Present address: Department of Physiology, University of Saskatchewan, Saskatoon, Canada.

One of the most marked actions of heparin is its reaction with proteins, first demonstrated by Fischer.[6] It is possible to demonstrate a reaction of heparin with almost any protein, provided a suitable method of testing is used.[7] With casein, as shown in FIGURE 1, the reaction can be demonstrated by the shift in the isoelectric point. Casein has an isoelectric point of 5.0. On the addition of heparin, the isoelectric point is shifted to 3.8. The reaction with protamine is of particular interest, since the fact that heparin has a greater affinity for protamine than for the plasma proteins allows the use of the former for the determination of heparin in blood. It has frequently been suggested that the anticoagulant action of heparin is an extension of this reaction with proteins to the proteins of the clotting system. While undoubtedly this action is involved, such a view ignores the unique anticoagulant action of heparin. This is shown when the anticoagulant activity of heparin on whole blood is compared with its activity on various isolated fractions of the clotting system. Thus, Jaques, Waters, and Charles,[8] isolated crystalline samples of heparin from different animal species and found marked differences in the specific anticoagulant potency of the samples (TABLE 1). However, these differences were much less marked when the activity was compared for antithrombic action or action on the clotting of blood plasma, while the reaction with protamine and the dye, Azure A, was found to be identical. This is further stressed by the recent studies of Astrup and Galsmar[9] who used an assay system for heparin similar to the well-known

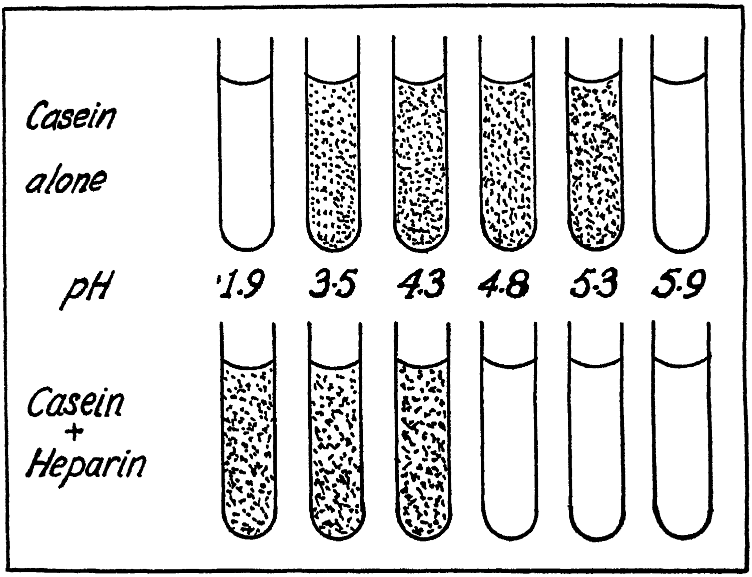


FIGURE 1. Effect of heparin on the isoelectric point of casein. 10 mg. of casein and 1 mg. of heparin in each 100 cc. of N/100 acetate buffer. (From Best).[43]

TABLE 1

COMPARISON OF THE RELATIVE ACTIVITIES OF HEPARIN FROM DIFFERENT SPECIES

All activities were determined by direct comparison with a standard beef heparin preparation of 100 units/mg. (From **Jaques, Waters & Charles**).[8]

Species	Protamine titer mg./mg. hep.	Meta-chromatic activity units/mg.	Antithrombin activity units/mg.	Anti-prothrombin activity units/mg.	Anti-coagulant activity units/mg.
Dog	2.1	92	140	250	240
Ox	1.9	92	100	100	100
Pig	2.2	106	69	60	44
Sheep	2.3	97	41	41	23

prothrombin time determination. In such a system, they found that the synthetic polysaccharide sulfuric acids (sulfonated cellulose and sulfonated starch) show greater anticoagulant activity than heparin, although their anticoagulant activity measured on fresh whole blood is only 1/20 that of heparin. As we have already indicated, such observations raise doubts as to the value of conclusions regarding the mechanism of the anticoagulant action of heparin based solely on studies of its antithrombic action.

Heparin *in vivo*. When the clotting time of blood is determined with increasing quantities of heparin added to the blood on removal, as shown in FIGURE 2, there is little change in the clotting time until a certain limiting concentration of heparin is reached, after which the clotting time

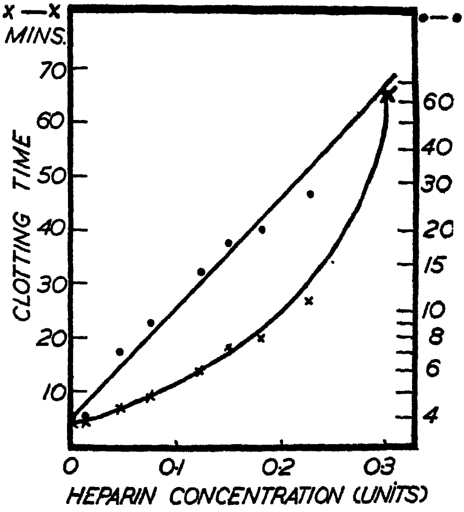
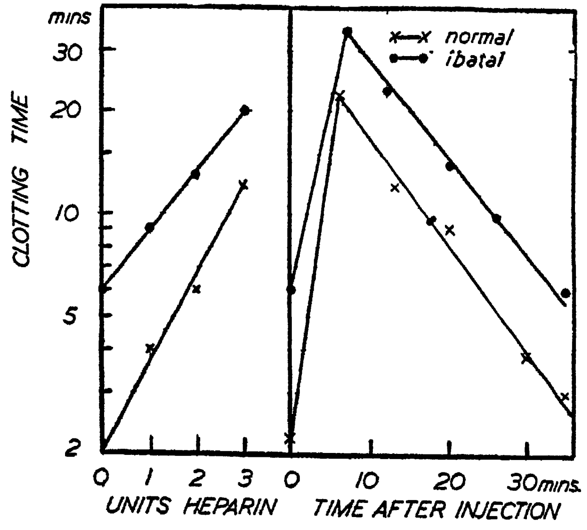


FIGURE 2. Relationship between clotting time and units of heparin added to blood *in vitro*. Blood samples taken from the exposed femoral vein of a dog under pentobarbital anaesthesia. Clotting times determined in the coagulometer of **Murray, Jaques, Perrett, & Best.**[13]

increases very rapidly. The relationship between clotting time and heparin concentration is logarithmic, so that a linear relationship may be obtained by plotting on semi-logarithmic paper.

On intravenous injection of heparin, the clotting time is prolonged proportionately to the dosage of heparin. However, the clotting time produced with moderate doses of heparin *in vivo* is considerably greater than that obtained on mixing the same quantity of heparin with the blood *in vitro*. This is due to the fact that some time is required for heparin to combine with the clotting system of the blood, in order for it to exert a maximal effect. This has been confirmed by suitable experiments *in vitro*. For the same reason, with moderate doses of heparin, the hypocoagulability of the blood after a single injection does not reach its peak until 5 to 10 minutes after the injection. The effect of a single intravenous injection of heparin is shown in FIGURE 3.



SEMI-LOG

FIGURE 3. The clotting time response to heparin *in vitro* and *in vivo*. Dog of 10 kgms. Clotting times determined in the coagulometer.[32] At zero time, 30 units/kgm. were injected intravenously. x—x, response of unanaesthetized animal. o—o, response after anaesthesia with sodium pentobarbital (ibatal).

The injection of 30 units/kgm. caused the clotting time to rise to a peak value of 34, five minutes after the injection. In thirty minutes, the hypocoagulability had disappeared and the clotting time returned to normal. The use of heparin has been suggested by de Takats[10] and others as a test for hypocoagulability or hypercoagulability of the blood and. also, as an indication of the need for anticoagulant therapy. The clotting time response to heparin *in vivo* gives us a heparin tolerance curve. Such a curve is due both to the state of coagulability of the blood itself, as well as to the ability of the body to dispose of heparin. For this reason, we have, in recent years, used a combined test in our experimental investigations. The clotting time response to four levels of heparin (0, 0.1, 0.2, 0.3 units/c.c.) added to the blood *in vitro* is first determined, and by the use of semi-log paper, a straight line is obtained. A single intravenous injection of 30 units/kgm. is then administered and the response *in vivo* measured. This is illustrated in FIGURE 3, in which the effect of anaesthesia (with pentobarbital) on the heparin response was tested. It can be seen that this caused a decreased heparin tolerance, which, however, was due to a decreased coagulability of the blood, since the change in the response to heparin *in vivo* paralleled the change *in vitro*.

As shown, the hypocoagulability due to a single intravenous injection of heparin wears off very quickly. What is the reason for this? There has been some debate as to whether heparin is excreted in the urine or not. Howell and Macdonald,[11] Wilander,[12] and Copley[13] have reported excretion, while Jaques,[14] and Reinert and Winterstein,[15] failed to find any excretion. Recent studies by our group and by P. Astrup[16] at Copenhagen, indicate the probable explanation for these differences. Both Astrup and Jaques find that with moderate doses of heparin (*i.e.*, at moderate blood levels), only a small portion of the heparin injected is excreted. The renal threshold for heparin was determined in the dog to be at a blood level of 1-2 units/ml.

At blood levels below the renal threshold, no excretion of heparin occurs. However, heparin disappears rapidly from the circulation, as seen in FIGURE 3. That this is not due to excretion is also shown by the fact that nephrectomy has no effect on this disappearance. We have recently observed the excretion of a degradation product of heparin, which we have termed uroheparin, when the blood concentration is below the threshold. Uroheparin gives a metachromatic reaction like heparin but has a much lower anticoagulant activity. The conversion of heparin to uroheparin is probably due to the enzyme, heparinase, previously described.[17] Astrup reports isolation of a substance similar to uroheparin from normal rabbit blood and from normal urine of both rabbits and human patients. In 25 patients studied, 4 showed an increase in the amount of this substance after operation.

Since heparin is both destroyed rapidly with increased concentrations and also shows a low renal threshold, the use of single large doses intravenously is a wasteful method of administration. Continuous intravenous administration was therefore recommended by the Toronto group when heparin was first introduced to clinical use. While we were not unmindful of the possibility of replacing this by a single intramuscular or subcutaneous injection, the possible dangers of a large quantity of heparin in a small localized area, seen experimentally, led us to avoid other types of administration. The use of subcutaneous and intramuscular injections of heparin has recently been developed. Walker[18] reports that a satisfactory hypocoagulability of prolonged duration can be obtained by intramuscular injection. In agreement with the experimental results of Jaques, Charles, and Best[19] he finds that subcutaneous injections give a somewhat erratic response. Loewe[20] has overcome this by the use of the Pitkin menstruum and has thus devised a subcutaneous method of administering heparin which apparently gives consistent results.

Determination of Heparin. Since this conference is on hemorrhage, the interest in heparin is probably focused on its contribution to larger problems. Therefore, a discussion of its determination will be of value.

The most satisfactory method is the isolation of the substance, preferably as the crystalline barium salt, by the Charles and Scott procedure, and its identification by its specific anticoagulant activity on fresh whole blood. This is the only method available for the determination of heparin in tissues. In this connection, attention should be drawn to the original observations of Charles and Scott[21] on the sensitivity of heparin to acid and alkali, and also to the greater specificity for heparin of the method of assay which uses fresh, whole blood. The metachromatic reaction of heparin with toluidine blue (or preferably Azure A) is of limited value. It is relatively insensitive, and proteins interfere with it. Since the reaction is also given by inactivated heparin, it may be of value in studying the process of inactivation. For the determination of heparin in blood, the effect on clotting time is the most useful method. The nature of the relationship between clotting time and heparin already referred to means that clotting times can be used to measure heparin concentration in the blood only over a very narrow range of concentrations. We have found, however, that by suitably modifying the method of determination, particularly in the procedure of taking the sample and the temperature of the determination, it is possible to standardize this method so that it is more or less sensitive to heparin. Taking the blood from an exposed vein or conducting the determination at low temperatures makes the determination sensitive to lower concentrations of heparin.

The titration of heparin with protamine may be a useful method. The least amount of protamine which can reduce the clotting time to normal is determined. Our initial communication[19] has been frequently quoted as establishing a conversion factor of 3:1 for mg. of protamine/mg. of heparin. However, it should be appreciated that the conversion factor changes with different lots of protamine. Also, the reaction shows a simple mass law relationship,[7] and hence, the conversion factor is changed if the conditions of the test are changed. Thus, the use of equal volumes of blood and protamine solution, as originally recommended, markedly increases the sensitivity of the determination.

A method of determination of heparin in blood has been developed by Astrup and Darling,[22] and by Volkert.[23] This is based on Astrup's accurate method for the determination of antithrombin. Applying this to rabbit blood, Volkert concludes that 20 per cent of the antithrombin present is heparin and that this constitutes the variable component. He finds that it is possible to cause the disappearance of the variable component, or heparin, by the injection of India ink.

Liberation of Heparin in the Body. One of the most important contributions to the study of heparin was the identification by Jorpes, Holmgren, and Wilander[24] of the mast cells as the heparin cells of the body. This localization of heparin in cells which are chiefly found in vessel walls, and in greatest numbers in tissues where marked disturbances to the blood by the ingress or egress of constituents is taking place (*e.g.*, lung, liver, intestine, thymus during involution; cf. Michels),[25] in itself strongly suggests that the physiological role of heparin in the body is to prevent clotting of the blood in vessels where such clotting is initiated by the products of various physiological processes. The evidence that clotting in the vessels will cause the secretion of heparin from the mast cells, has been provided by study of the clotting system in anaphylactic and peptone shock.

In 1909, Biedl and Kraus[26] first reported increased clotting time of the blood of dogs in anaphylactic shock. In 1940, Jaques and Waters[27] reported the quantitative isolation of heparin from the blood of dogs in anaphylactic shock; Wilander,[12] in 1938, had reported its separation from the blood of dogs in peptone shock. Since a definite response to the antigen could be obtained in the absence of the liver while no liberation of heparin occurred, it is evident that the heparin came from the liver. This was also demonstrated by determinations of the heparin content of the liver after shock. These results on heparin parallel the observations of others on histamine, which is likewise released from the liver in anaphylactic shock in the dog.

Regarding the mechanism that causes the liberation of heparin from the liver in anaphylaxis and peptone shock, it might be presumed that this is simply the anaphylactic response of the mast cells. However, Rocha e Silva and Teixeira[28] obtained indications that it depends on

TABLE 2

QUANTITY OF HISTAMINE AND HEPARIN LIBERATED BY PEPTONE FROM THE ISOLATED DOG'S LIVER

(From **Rocha e Silva, Scroggie, Fidlar, & Jaques**).[30]

Dog No.	Liver gm.	Perfusion fluid	Histamine μg	Heparin mg.
11	170	Tyrode's solution	26.8	none
17	402	Tyrode's solution	38.0	traces
15	140	Blood in silicone	417.7	5.45
14	359	Blood in silicone	607.7	8.15
12	—	Blood in silicone	3313.	16.5
20	470	Blood in silicone	8550.	25.35
17	402	Heparinized blood	93.7	1.75
19	365	Defibrinated blood	1162.	not estimated

changes induced in the blood. This has been further demonstrated by the experiments of Rocha e Silva, Scroggie, Fidlar, and Jaques[29] on the isolated dog liver. When peptone is added to Tyrode's solution perfused through the isolated liver, no heparin and only a trace of histamine is liberated. When citrated blood and peptone are perfused, only small variable amounts of histamine and heparin are liberated (TABLE 2). Recently, Jaques, Fidlar, Feldsted, and Macdonald[30] have described a technique using one of the silicone plastics, whereby blood can be preserved from clotting for several hours without adding any anticoagulant. When peptone is added to such blood before perfusion through the liver, enormous amounts of heparin and histamine are liberated. If, however, clotting of the blood is prevented by the addition of heparin to it, then the addition of peptone may not be effective in liberating heparin from the liver. As with the action of heparin on clotting (cf. above), and on thrombosis,[31] this inhibitory action of heparin is most marked if the heparin remains in contact with the blood for 45 minutes before testing the blood. These findings suggest that the changes in the blood which heparin prevents are likewise those which cause its secretion into the blood.

Heparin and Thrombosis. It has long been debated whether thrombosis is the same as intravascular clotting. The appearance of the so-called white thrombus as an almost pure mass of platelets has led many observers to make a marked distinction between a thrombus and a clot, with the former defined as an agglutination of platelets. However, Murray, Jaques, Perrett, and Best,[32] and also Best, Cowan, and MacLean[33] were able to show that heparin could prevent thrombosis produced experimentally. The recent clinical data on the use of heparin have completely confirmed the original experimental findings. As a result of this, it was postulated that heparin may act independently to prevent platelet agglutination. However, as has been shown by Dale and Jaques,[34] Richards and Cortell,[35] and Bollman and Preston,[36] dicumarol, which lowers the prothrombin level of the blood, is also effective in preventing thrombosis. Wright,[37] and Spooner and Meyer[38] have shown that platelet adhesiveness is reduced at the same time. While these findings suggest that heparin and dicumarol act to prevent thrombosis by virtue of their anticoagulant action, a direct effect on platelets is not ruled out.

In this connection, it is of interest, as first shown by Copley and Robb,[39] that many commercial preparations of heparin will cause agglutination of platelets *in vitro* and *in vivo* (TABLE 3). This phenomenon has not yet been studied with heparin made from the same species. Piper[40] has reported a similar effect with the synthetic polysaccharide sulfuric acids, which is not surprising when it is appreciated that intravenous injection of many colloids (cf. Roskam),[41] *e.g.*, gelatin and glycogen, will cause platelet agglutination. The agglutination of the platelets by these substances on injection results in a marked fall in the platelet count due to

TABLE 3

EFFECT OF INTRAVENOUS INJECTIONS OF HEPARIN ON THE PLATELET COUNT*

Experiment	Heparin units/kg.	Minutes after heparin	R.B C x 10^3/mm.^{3}	W.B.C. /mm.^{3}	Platelets x 10^3/mm.^{3}	Per cent agglutination
22	500	—	5,588	12,100	495	0
		3½	5,594	12,200	135	56
3	500	—	4,763	9,400	336	0
		2½	4,956	8,000	143	59
		15	4,781	2,600	148	12
24	100	—	5,794	10,000	295	1
		5	5,781	5,200	20	24
		37	5,981	9,100	228	5
25	1500	—	5,369	8,500	341	0
		3½	5,363	1,800	15	27
		60	6,013	3,800	115	42
7	50	—	6,306	9,800	397	0
		4	6,344	3,800	39	48
		65	5,675	3,800	234	27
		407	6,881	14,000	203	21
	17¾ hours after†		6,400	25,000	191	0

* We are indebted to Dr. E. Fidlar for these unpublished data.
† Continuous injection of 1 unit/kg /min.

the filtering out of the clumps of platelets in the capillaries. However, there are no symptoms arising from this clumping, and the clumps evidently disintegrate and return the platelets to the blood, since in one or two hours the count is back to normal. In such clumps, the individual platelet can still be identified. Evidently, the agglutination of platelets itself is not a serious phenomenon and may be a physiological process which occurs frequently. If agglutination proceeds to the stage of disintegration with the release of various physiologically active substances, then formation of a thrombus will occur. We would suggest that it is at this stage that heparin and dicumarol act to prevent thrombosis. Rocha e Silva[42] has suggested that these same changes in platelets and leucocytes are necessary factors in the secretion of heparin by the mast cells under the influence of peptone. Until the effect of homologous heparin on the agglutination of platelets has been studied, the suggestion that the agglutination of platelets by heparin is of physiological significance in the repair process of slightly damaged blood vessels is not warranted. It is just possible, however, that heparin may be proved to have this action and, consequently, to exert the secondary effect of preventing the extension of the physiological process to a pathological one, *i.e.*, to gross thrombus formation. It is also possible that the agglutination of platelets by many preparations of heparin is due to an impurity, and that this will eventually be separated from heparin. Investigation of this problem is proceeding in our laboratories.

BIBLIOGRAPHY

1. **Brinkhous, K. M., H. P. Smith, E. D. Warner, & W. H. Seegers**
   1939. Am. J. Physiol. **125**: 683.
2. **Astrup, T., & S. Darling**
   1943. Acta Physiol. Scand. **5**: 13.
3. **Feisly, R., & M. Enowicz**
   1944. J. Suisse Méd. **13**: 274.
4. **Horwitt, M. K.**
   1944. J. Biol. Chem. **156**: 427.
5. **Rocha e Silva, M., & S. O. Andrade**
   1945. Science **92**: 670.
6. **Fischer, A.**
   1935. Biochem. Z. **278**: 133.
7. **Jaques, L. B.**
   1943. Biochem. J. **37**: 189.
8. **Jaques, L. B., E. T. Waters, & A. F. Charles**
   1942. J. Biol. Chem. **144**: 229.
9. **Astrup, T. & I. Galsmar**
   1944. Acta Physiol. Scand. **8**: 361.
10. **De Takats, G.**
    1943. Surg., Gynec. & Obstet. **77**: 31.
11. **Howell, W. H., & C. H. McDonald**
    1930. Bull. Johns Hopkins Hosp. **46**: 365.
12. **Wilander, O.**
    1938. Skand. Arch. Physiol. **81** (Suppl. 15).
13. **Copley, A. L.**
    1941. Science **93**: 478.
14. **Jaques, L. B.**
    1939. Am. J. Physiol. **125**: 98.
15. **Reinert, M., & A. Winterstein**
    1939. Arch. Internat. Pharmacodyn. & Therap. **62**: 47.
16. **Astrup, P.**
    1944. Heparinundersogelser. Copenhagen.
17. **Jaques, L. B.**
    1940. J. Biol. Chem. **133**: 445.
18. **Walker, J.**
    1945. Surgery **17**: 54.
19. **Jaques, L. B., A. F. Charles, & C. H. Best**
    1938. Acta Med. Scand. Suppl. **90**: 190.
20. **Loewe, L., P. Rosenblatt, & M. Lederer**
    1942. Proc. Soc. Exp. Biol. & Med. **50**: 53.
21. **Charles, A. F., & D. A. Scott**
    1933. J. Biol. Chem. **102**: 437.
22. **Astrup, T., & S. Darling**
    1942. Acta Physiol. Scand. **4**: 293.
23. **Volkert, M.**
    1942. Acta Physiol. Scand. **5** (Suppl. 15).
24. **Jorpes, E., H. Holmgren, & O. Wilander**
    1937. Z. mikro-anat. Forsch. **42**: 279.
25. **Michels, N. A.**
    1938. The Mast Cells. In: *Downey's Handbook of Haematology* **1**: 234.
26. **Biedl, A., & R. Kraus**
    1909. Wien. klin. Wochenschr. **22**: 363.
27. **Jaques, L. B., & E. T. Waters**
    1941. J. Physiol. **99**: 454.
28. **Rocha e Silva, M., & R. M. Teixeira**
    1946. Proc. Soc. Exp. Biol. & Med. **61**: 376.

29. **Rocha e Silva, M., A. E. Scroggie, E. Fidlar, & L. B. Jaques**
1947. Proc. Soc. Exp. Biol. & Med. **64**: 141.
30. **Jaques, L. B., E. Fidlar, E. T. Feldsted, & A. G. Macdonald**
1946. Can. Med. Assoc. J. **55**: 26.
31. **Solandt, D. Y., & C. H. Best**
1940. Lancet. **1**: 1042.
32. **Murray, D. W. G., L. B. Jaques, T. S. Perrett, & C. H. Best**
1937. Surgery **2**: 163.
33. **Best, C. H., C. Cowan, & D. L. MacLean**
1938. J. Physiol. **92**: 20.
34. **Dale, D. U., & L. B. Jaques**
1942. Can. Med. Assoc. J. **46**: 546.
35. **Richards, R. K., & R. Cortell**
1942. Proc. Soc. Exp. Biol. & Med. **50**: 237.
36. **Bollman, J. L., & F. W. Preston**
1942. J. A. M. A. **120**: 1021.
37. **Wright, H. P.**
1945. J. Path. & Bact. **57**: 382.
38. **Spooner, M., & O. O. Meyer**
1944. Am. J. Physiol. **142**: 279.
39. **Copley, A. L., & T. P. Robb**
1942. Am. J. Clin. Path. **12**: 416.
40. **Piper, J.**
1945. Acta Physiol. Scand. **9**: 28.
41. **Roskam, J.**
1927. Physiologie normale et pathologique du globulin. Paris.
42. **Rocha e Silva, M.**
Histamina e anafilaxia. Edigraf. Sao Paulo, Brazil.
43. **Best, C. H.**
1940. Harvey Lectures **36**: 66.

DISCUSSION OF THE PAPER

Dr. Marjorie B. Zucker *(Department of Physiology, College of Physicians and Surgeons, Columbia University, New York, N. Y.):*

I should like to report some experiments which have been published recently in the American Journal of Physiology. Spontaneous hemostasis was studied in rats under nembutal anesthesia, by microscopic observation of the mesoappendiceal venules and of the large branches of the superior mesenteric artery and vein which lie in the mesentery. When the muscular mesenteric vessels are nicked, vasoconstriction of the injured vessel and of the adjacent, parallel, uninjured vessel occurs near the injury. A plug of blood platelets, chiefly extravascular, forms at the site of hemorrhage. Fibrin is not histologically demonstrable in this plug. Bleeding ceases in about 2 minutes, but frequently recurs for brief intervals, perhaps as a result of the experimental conditions which prevent the formation of a fibrin clot in the wound. The hemostatic response of rats after abdominal sympathectomy does not differ from that of normal animals.

Non-muscular venules in the mesoappendix do not contract following transection. When a dull blade is used, hemorrhage is usually entirely prevented, probably by closure of the vessel stumps by adherence of the endothelium, by pressure of connective tissue fibers drawn over the cut vessel, or by stasis, which frequently occurs. Transection with a sharp blade is followed by bleeding which is arrested by the formation of a platelet plug.

Rats with thrombocytopenic purpura induced by the injection of anti-platelet serum fail to form platelet plugs, since the remaining circulating platelets are incapable of forming thrombi. Bleeding from non-muscular venules or from large arteries or veins is not arrested. In the mesentery, the nicked vessels contract, but no contraction of the adjacent uninjured vessels occurs. This strongly suggests that in normal animals, this contraction is induced by vasoconstrictor substances liberated from the platelet plug.

In some rats, 30 to 45 minutes after the intravenous administration of large doses of heparin (500 to 2250 units/kg.), no platelet plug forms at the site of injury to the mesenteric vessels, and bleeding is incessant. In other heparinized animals, and in rats in which the prothrombin level is markedly lowered by the administration of dicumarol, plugs are produced which are ineffective in arresting bleeding. The incised vessel contracts normally, but the uninjured vessel contracts only in the presence of a platelet plug. Experiments reported by Copley and Lalich suggest a difference in the mechanism of action of heparin and dicumarol on the bleeding time and clot resistance of mice, which was not apparent in my experiments.

In conclusion, the experiments indicate that non-muscular venules do not contract when cut. When bleeding occurs, hemostasis is produced by the formation of a platelet plug, or bleeding may be entirely prevented by other factors. In larger vessels, hemostasis is produced mainly by the formation of a platelet plug. The contraction of nicked muscular blood vessels, which is apparently largely the result of mechanical stimulation of smooth muscle, is incapable in itself of arresting hemorrhage. The vasoconstrictor substances released from the platelet plug doubtless contribute to this contraction and, in addition, produce constriction of adjacent uninjured blood vessels.

Dr. Sylvan E. Moolten *(St. Peter's General Hospital, New Brunswick, N. J.):*

EXTRACTABLE FACTORS IN THE SPLEEN AND OTHER ORGANIC SOURCES WHICH INFLUENCE THE BLOOD PLATELET COUNT*

In connection with the foregoing papers dealing with a number of tissue factors concerned with the clotting mechanism and the behavior of blood platelets in circulation, it may be appropriate to discuss two lipoid substances which have powerful but mutually antagonistic effects on the platelet count.

The existence of a hormone-like element in the spleen which restricted the rate of platelet production by megakaryocytes of the bone marrow was surmised as early as 30 years ago by Frank,[1] who termed it a "myelotoxic influence". Torrioli[2] and, more recently, Hobson and Witts[3] reported thrombocytopenic activity in watery extracts of normal spleen. Troland and Lee[4] (1938) employed filtered watery suspensions of the gummy residue obtained when acetone extracts were evaporated to dryness. The spleens of three patients with thrombocytopenic purpura were assayed as well as other types of spleen, fibromyomatous uterus, and thyroid tissue. By intravenous injection of the extracts from the spleens of purpura, the authors obtained marked thrombocytopenia in rabbits and other animals, but negative results otherwise. Numerous other observers employing acetone extraction attempted to repeat their results and reported conflicting findings.[3, 5] Positive results were obtained with extracts of normal spleen and the spleen of various abnormal states besides purpura, and many negative results were reported. One observer reported a rise in platelet count in one of several experiments.[6]

---

* This paper is also a discussion of the paper by Dr. Ferguson (see pages 486-500).

The present observations, reported in 1945,[7] originated the year before, when the spleen of a typical case of thrombocytopenic purpura was secured by operation. The histology of the spleen was noteworthy for the widespread hyperplasia of reticulum cells, especially about malpighian follicles where they also formed discrete nodules of epithelioid cells. Because of the presence of vacuoles in their cytoplasm, suggesting fatty substances, frozen sections were stained with Sudan IV and showed weakly staining sudanophilic globules.

Part of the spleen was subjected to acetone extraction. Four-fifths of the brownish gummy residue after acetone removal was shaken with distilled water and filtered. The milky filtrate (Troland-Lee extract), injected intravenously, caused a slight rise in platelets, followed by a fall which was maintained by successive doses until the fourth day, when the rabbit died. The filter-paper residue (discarded in experiments by previous observers) was given in peanut oil to another rabbit in successive intramuscular injections. The fall in platelet count was much more marked and sustained, but the count rapidly returned to normal when the injections were ended. The remaining fifth of the original acetone residue was extracted with ether and given in peanut oil to another rabbit in a single dose. An immediate thrombocytopenic effect was pronounced, and thrombocytopenia persisted in diminishing degree for several weeks. The ether-insoluble, water-soluble residue was without effect.

Having satisfied ourselves that the active substance, previously named "thrombocytopen" by Troland and Lee, was associated with the cell lipoids of this spleen, we sought evidence for its existence in normal spleen. Thrombocytopen was detected, though in smaller amounts than in purpura. The active substance was found to produce its quickest and most consistent response when given in aqueous suspension intramuscularly. Several rabbits were standardized with equivalent doses.

A discordant finding was encountered in the spleen of a case of Hodgkin's disease. The acetone-soluble material, given in the manner described, produced, instead of a drop, a sharp rise in platelet count, which was duplicated on subsequent trial. By means of several methods of fractionation, including fractional crystallization from petroleum ether and from methanol, it was established that this spleen contained not only much thrombocytopen but an even greater amount of a lipoid which had the opposite effect on the platelet count, *viz.,* the production of marked thrombocytosis. The latter substance was therefore called "thrombocytosin". It differed from thrombocytopen in its greater ease of solution in fat-solvents and its greater ease of dispersion in water. Additional refinements in the procedure of extraction and separation made it possible for us to obtain consistent, large effects with doses of 5 mg. of either factor. It remained for Dr. Harry Sobotka, however, to uncover the important fact that, even with such small doses as 5 mg., the largest amount of the material injected was pure cholesterol in finely dispersed form, serving as a vehicle for extremely minute amounts of the active factors. The latter were subsequently purified from cholesterol and were found to be oily substances with fishy odor, non-crystallizable, and forming insoluble digitonides in the manner of steroids.

The separation procedure now developed made it possible to assay a number of tissues and other substances for thrombocytopen and thrombocytosin. Both of these were present in normal spleen, human and bovine, and could be found in reduced amounts in fibrotic and leukemic spleens in inverse proportion to the amount of metaplasia or replacement. Brain tissue, liver, and defatted myocardium contained insignificant amounts of either factor. Lymph nodes contained considerable thrombocytosin but little or no thrombocytopen. Subcutaneous fat gave very high yields of thrombocytosin but was apparently devoid of thrombocytopen. Perirenal fat was inert. Egg yolk fluctuated in yield of thrombocytopen but gave high yields of thrombocytosin. Normal urine contained both factors: the urine of a splenectomized individual contained no thrombocytopen and about 15 times the normal maximum of thrombocytosin. Vegetable oils were negative for either factor.

Administration of large doses of thrombocytopen (obtained from beef spleen) to a rabbit over a period of several days maintained a low platelet count (without purpura) until the eleventh day, when spontaneous "escape" occurred. As the platelet count reached the pre-injection level, the animal was sacrificed. Its bone marrow exhibited striking increase in the number of megakaryocytes which, for the most part, were immature. This finding recalls the observation of Limarzi[8] in human idiopathic thrombocytopenic purpura before splenectomy, and probably represents a type of maturation arrest under the influence of thrombocytopen.

The results obtained thus far in these studies suggest a physiological role of the spleen as the source of a steroid hormone, thrombocytopen, normally concerned in balancing the thrombocytosis-producing effect of a fat-soluble dietary factor, thrombocytosin. Purpura hemorrhagica, when associated with reticulum cell hyperplasia of the spleen, may be regarded as the result of an over-production of thrombocytopen because of splenic hyper-function. In some cases (pubertal metrorrhagia,[9] pregnancy[10]), this may be part of an endocrine derangement.

Because of change of army station, it was impossible for the author to complete many further details of this work during active military service. The work is just now being resumed with special attention to the role of subcutaneous trauma, accidental or surgical, in the mobilization of stored thrombocytosin and the initiation of postoperative thrombosis. Other problems to be undertaken shortly will include differential assay of urine of patients with thrombocytopenic purpura complicating such conditions as pregnancy and infectious mononucleosis, and of various patients with other abnormalities of the platelet count and clotting mechanism. Therapeutic trials are also to be conducted with each factor, in suitable conditions. Further work is contemplated, too, on methods for improved separation and assay.

References

1. **Frank, E.**
   1915. Berl klin. Wochenschr. **52**: 454, 490.
2. **Torrioli, M., & V. Puddu**
   1938. J. A. M. A. **111**: 1455.
3. **Hobson, F. C. G., & L. J. Witts**
   1940. Brit. Med. J. **1**: 50.
4. **Troland, E., & F. Lee**
   1938. Bull. Johns Hopkins Hosp **62**: 85.
   1938. J. A. M. A. **111**: 221.
5. **Rose & Boyer, Otenasek & Lee, Paul, Uihlein, Cronkite, Pohle & Meyer, Hodge & Strong, Colmer & Mersheimer, Moore, Major & Weber, and Tocantins**
   Cited by **Moolten.**[7]
6. **Major & Weber**
   Cited by **Moolten.**[7]
7. **Moolten, S. E.**
   1945. Studies on extractable factors in the spleen and other organic sources which influence the blood platelet count. J. Mount Sinai Hosp. **12**: 866.
8. **Limarzi, L. R., & E. M. Schleicher**
   1940. J. A. M. A. **114**: 12.
9. **Goldburgh, H. L., & B. A. Gouley**
   1940. Am. J. Med. Sci. **200**: 499.
10. **Polowe, D.**
   1944. J. A. M. A. **124**: 771.

Dr. A. L. Copley *(New York University, New York, N. Y.)*:

Our findings that heparin does not prevent the agglutination of platelets[1] have been further substantiated by observations on the whole blood of different animal species and by agglutination tests with heparinized plasma of man, dog, and swine.[2] Heparin alone has no effect on the agglutination of isolated platelets. As already noted previously,[1] a plasma factor is needed to bring about platelet agglutination. In this connection, our observations on *Limulus* blood are of interest. Sodium heparin *in vitro* and *in vivo* in excessive amounts does not inhibit the agglutination of amoebocytes or the subsequent gelation mentioned in the discussion of Dr. Ferguson's paper.

The contention that platelet agglutination and fibrin formation are governed by the two separate mechanisms has been revived by Dr. Best and his associates.[3, 4] They believe that an "adhesive agent" is needed for platelet agglutination. Dr. Houlihan and I have shown that the agglutination of platelets is not brought about by fibrin formation and that the two processes have different mechanisms. The "adhesive agent" or the platelet-agglutinant substance appears to reside in the globulin fractions.[5] We found that a platelet suspension from one subject showed significant variations in the degree of agglutination when tested with plasma and sera from different subjects.

We should differentiate between adhesiveness of platelets to the vessel wall and agglutination of platelets.[2] There may be also a difference between the adhesiveness of platelet agglutinates and that of the endothelial lining which, under certain conditions, according to Chambers and Zweifach, may exhibit adhesive properties. Dr. Houlihan and I showed that platelet agglutinates are not necessarily adhesive. We have verified Helen Wright's observations on platelet counts in rotating heparinized blood samples, although we disagree with her interpretations. We found that human or dog platelets which were incubated with heparin plasma have two adhesive properties, *viz.,* high degree of agglutinability, and also adhesiveness to the glass wall. In studying coagulation thrombi in segments of artery and vein in dogs, we demonstrated that blood coagula are not adhesive to the living, non-infected wall of the blood-vessel.[7]

It may be also necessary to distinguish between coagulation thrombi, agglutination thrombi, and mixed thrombi. In the latter, there are both blood coagulation and agglutination of platelets. Heparin is useful in the treatment of thrombosis because of its anticoagulant effect. Since we have shown that prothrombin contains a platelet-agglutinant factor, it remains to determine whether hypoprothrombinemia induced by dicumarol also lowers the platelet-agglutinability of such plasma. Anticoagulants *per se* may not be of therapeutic value in certain clinical conditions initiated by platelet thrombi.

I do not question the results of Dr. Best and his associates[4] that large doses of heparin *in vivo* may prevent platelet agglutination. I am merely trying to find a reason for the discrepancy between *in vivo* and *in vitro* findings. This might be explained as follows. We studied the anomalous flow properties of blood in a rolling ball viscometer[7] and found that heparin in increasing amounts tended to decrease the yield value, apparent viscosity, and pseudoviscosity of blood. I wonder whether this tendency of increased blood fluidity with large doses of heparin may explain the findings of Best *et al.* Thus, rapid blood flow past the injured area may prevent agglutination of platelets.

I should like to ask Dr. Best whether he made any platelet counts and determined whether there may be seen platelet clumps in blood samples obtained following heparinization.

References

1. **Copley, A. L., & T. P. Robb**
   1942. Am. J. Clin. Path. **12**: 416, 536.
2. **Copley, A. L., & R. B. Houlihan**
   1947. Blood. (In press.)
   Houlihan, R. B., & A. L. Copley
   1946. J. Bact. **52**: 439.
3. **Best, C. H.**
   1940-41. Harvey Lectures. Ser. **36**: 66.
4. **Best, C. H., C. Cowan, & D. L. MacLean**
   1938. J. Physiol. **92**: 20.
5. **Copley, R. L., & R. B. Houlihan**
   1945. Fed. Proc. **4**: 173.
6. **Copley, A. L., & P. L. Stefko**
   1947. Surg., Gynec. & Obst. (In press.)
7. **Copley, A. L., L. C. Krchma, & M. E. Whitney**
   1942. J. Gen. Physiol. **26**:49.

Dr. L. B. Jaques:

In presenting our paper, we were, of course, aware of Dr. Zucker's important contributions concerning the significance of the platelets in hemostasis and the action of heparin on this. Time, however, did not permit us to include an appreciation of her important findings that would relate them to our own work. Several points in Dr. Ferguson's paper reminded me of the common observation that the rate of formation of a clot appears to affect its strength. I would like to ask Dr. Zucker if the slow rate of formation of the plugs following large doses of heparin is a factor in the poor hemostasis.

With regard to Dr. Moolten's paper, I wish to express my admiration for a fine piece of work on a problem which has long demanded attention. We have conducted a few preliminary studies along this line in the last few months. We have found that aqueous extracts of rat spleen will cause a definite increase in the platelet count in splenectomized rats, lasting for 6-8 hours. However, with our crude extracts, we found no increase in the platelet count in rats that were not splenectomized. I would like to ask Dr. Moolten if he has studied the action of thrombocytopen and thrombocytosin in the splenectomized animal, and in species other than the rabbit.

With regard to Dr. Copley's point on heparin and platelet agglutination, we can merely repeat that it has been demonstrated frequently and conclusively that heparin prevents the formation of a thrombus experimentally and the phenomenon of thrombosis clinically, while we are in complete agreement with him that heparin will agglutinate platelets.

In concluding this discussion on heparin, I would like to make the point that, due to its ability to react with proteins (Biochem. J. **37**: 189), heparin may cause many non-specific effects when present in high concentrations. Caution must be exercised in making generalizations, since the concentrations required for study of the problems discussed are in this range.

THE RELATIONSHIP OF VITAMIN K TO HEMORRHAGE AND COAGULATION

By S. A. THAYER
Department of Biochemistry, St. Louis University School of Medicine,
St. Louis, Missouri

Prior to 1880, hemorrhage in obstructive jaundice and allied conditions was merely an interesting observation that had little practical importance. With the introduction of biliary surgery, however, bleeding took on a new significance.

In the early days of gall bladder surgery, the mortality from hemorrhage, although high, was not particularly striking, since bleeding was only one of the many causes of death. As surgical technique, pre- and postoperative care improved, the percentage mortality from hemorrhage steadily increased, since death from other causes was greatly reduced. Actually, the true incidence of cholemic bleeding was also decreased, especially after the introduction of the therapeutic practice of giving glucose intravenously. The work of Ravdin and his associates, who found glucose more effective than calcium in the treatment of cholemic bleeding, is now understandable. Nevertheless, in spite of all efforts, hemorrhage remained one of the major hazards of biliary surgery and, according to Butt and Snell, imposed a risk of death of approximately 5 per cent.

Earlier attempts to find the cause of the bleeding in obstructive jaundice were unsuccessful. Many of the earlier investigators believed that a deficiency in fibrinogen was the cause, but later studies failed to substantiate this view. Calcium was reported, by various workers, to be the causative factor and the use of calcium was recommended for the treatment of cholemic bleeding. Ultimately, sufficient evidence accumulated to overthrow the calcium deficiency theory. The presence of bile pigments and salts was also considered as a probable cause of the bleeding tendency. Anticoagulants were postulated as factors responsible for hemorrhagic diathesis. Just prior to the entry of vitamin K into medicine, Carr and Foote postulated that sulfur amino acids and related compounds accumulated in the blood of jaundice patients and interfered with coagulation. This theory enjoyed considerable popularity.

The beginning of the solution of cholemic bleeding dates back to 1907, when Morawitz and Bierich found that the delayed clotting in jaundiced blood was due to a deficiency of fibrin ferment. They erroneously attributed this defect to a lack of thromboplastin. Six years later, Whipple, using a method essentially the same in principle to that employed later by Quick for the quantitative determination of prothrombin, found that the clotting defect in jaundice was due to diminished prothrombin.

Bancroft and his associates, in 1929, applied the coagulation test of recalcified oxalated plasma to various clinical conditions. These investigators, confirming earlier work, found that the clotting time was definitely delayed in hemophilia and in certain cases of jaundice.

Quick, in 1935, developed a quantitative method for the determination of prothrombin. By means of this test, the important finding was made that a marked decrease in the prothrombin of the blood occurred in certain cases of obstructive jaundice, which led him to conclude that the hemorrhagic tendency was probably caused by a lack of this clotting factor.

In 1929, Dam, incidental to a study of cholesterol metabolism, observed that chicks on a synthetic diet developed a hemorrhagic condition which was neither prevented nor cured by citrus juice. In 1931, McFarlane, Graham, and Richardson observed a similar bleeding condition in chicks fed fish meal extracted with ether. They found that vegetable proteins prevented the bleeding condition while cod liver oil was ineffective. They also noted that the blood failed to clot.

After a lapse of two years, Holst and Halbrook again described this bleeding disease and stated that it was cured by feeding cabbage. They concluded that it was caused by a vitamin C deficiency in spite of earlier findings to the contrary. In 1934, Dam and associates resumed the study of this hemorrhagic chick disease. They found that the bleeding tendency could be prevented by feeding cereals or seeds, and they again demonstrated that vitamin C had no curative action. In 1935, Dam concluded that the bleeding was due to a deficiency of a hitherto unrecognized substance which was different from any of the known vitamins. He gave it the name "Vitamin K" (*Koagulationsvitamin*). At the same time, work on this chick disease was continuing at the University of California. Halbrook found that 5 per cent of dehydrated alfalfa, or an equivalent amount of an ether extract of this material, prevented the appearance of the hemorrhagic condition. Almquist and Stokstad then entered the field and in their first publication completely verified Dam's findings concerning this new vitamin and contributed much additional information.

Schonheyder made the next important contribution. Early in his work, he suspected that a deficiency of prothrombin occurred. He reported, in 1936, that the most probable explanation of the prolonged clotting time in chicks suffering from vitamin K deficiency is due to reduced concentration of prothrombin. Quick also found: (1) that the prothrombin rapidly decreased in chicks fed a vitamin K-free diet; (2) that the bleeding tendency appeared when the prothrombin dropped to about 10 per cent of normal; and (3) that alfalfa promptly and effectively restored the prothrombin.

Source of Vitamin K. A survey of sources of vitamin K by Dam in Copenhagen, and by Almquist in California, showed that this new vitamin was found chiefly in green leaves. Spinach and alfalfa are rich sources. Thayer *et al.* showed that dried alfalfa leaf meal contains about 10 micrograms per gram or one part in 10,000, *i.e.*, the weight of alfalfa is only 0.01 per cent vitamin K.

Almquist observed another important source, namely, the production of vitamin K by microorganisms. This was utilized by the Mayo workers in the production of their first product for clinical testing, an observation which we extended in the study of putrefied fish meal from which a second natural vitamin K was isolated.

Bioassay. Little progress could be made until suitable methods for determining vitamin K activity were available. The first methods were crude and theoretically faulty, but they furnished amazingly correct results. The first step in this study was the development of a diet which was practically free of vitamin K, but otherwise essentially complete.

Almquist devised a simple but effective ration which has remained the standard for most of the research carried out in this country.

Dam also devised a diet which is somewhat more complex but satisfactory. Ansbacher has described another diet in which traces of vitamin K are destroyed by prolonged heating. According to Almquist, this ration lacks other essential factors, particularly those needed for growth.

The second step in the assay was the establishment of criteria of vitamin K deficiency. The bioassay depends on the capacity of the product under study to restore to normal the clotting time of chicks which have been maintained on an ether-extracted diet. After a period of about two weeks, the clotting time of chicks on the deficient diet will exceed two hours. This is due to a deficiency of prothrombin in the blood. The administration of 1 microgram of vitamin K will cause the appearance of more prothrombin in the blood and the clotting time is restored to normal, *i.e.*, $>$ 10 minutes, within less than six hours.

Almquist and Klose, in 1939, adopted Quick's prothrombin time method and made it the basis of their assay method. Tidrick, Joyce, and Smith recently have employed the two-stage method for determining prothrombin as an assay method.

In spite of the objections to the clotting time as a criterion of vitamin K deficiency, a number of investigators including Dann, Ansbacher, and our own group at St. Louis University have successfully developed and employed assay methods based on this determination.

The chick is used universally because of the difficulty in producing prothrombin deficiency in mammals. A mild deficiency has been produced by Dam and Glavind in a small proportion of rats by dietary means alone. Deficiency can be produced in rats by a bile fistula to prevent bile reaching the intestine, as demonstrated by Greaves. Elliot *et al.* have shown that rats fed on a diet containing mineral oil (20 per cent) soon showed a marked lowering of prothrombin. Both of these procedures prevent absorption of vitamin K.

At this point, it is desirable to discuss briefly the use of prothrombin time rather than clotting time. It has been found that clotting time in obstructive jaundice may be within normal limits in cases in which the concentration of prothrombin is only 25 per cent of the normal value. Consequently, the danger of impaired coagulation is not revealed by the gross clotting time, but is detected by the more delicate determination of prothrombin.

The gross features of the clotting mechanism are:

$$\text{Prothrombin} + \text{Thromboplastin} + \text{Ca}^{++} \xrightarrow{\qquad} \text{Thrombin}$$
$$\text{Thrombin} + \text{Fibrinogen} \xrightarrow{\qquad} \text{Fibrin (clot)}$$

The Quick method provides optimal amounts of added thromboplastin and calcium, so that the conversion of prothrombin to thrombin is exceedingly rapid. The rate of clot formation is then dependent on the amount of thrombin. Normal blood gives a prothrombin time of about twenty seconds, but with blood from patients with obstructive jaundice the prothrombin time may be over two minutes.

Dam and his associates, and Warner *et al.* introduced different procedures for the determination of prothrombin. The literature abounds with papers describing modifications of the three methods.

Vitamin K in Treatment of Disease:

OBSTRUCTIVE JAUNDICE. The most important cause of hypoprothrombinemia in the adult is obstructive jaundice. The absence of bile salts in the intestines is responsible for inadequate absorption of the fat-soluble vitamin K to meet the physiological needs. It is a fallacy, however, to suppose that every jaundiced patient will bleed. Nevertheless, it is fairly safe to suspect a deficiency of prothrombin in a patient with a complete and prolonged obstructive jaundice. Vitamin K, or the water-soluble antihemorrhagic compounds, are effective therapeutic agents in restoring the prothrombin content of the blood. These water-soluble compounds are superior to the natural water-insoluble vitamin K in that they are absorbed from the intestinal tract in the absence of bile and, also, that on intravenous administration they cause prompt appearance of an increased quantity of prothrombin in the blood. This has been demonstrated for salts of the phosphate, sulfate, succinate, etc., by Warner and Flynn, Smith and Owen, Foster *et al.*, Almquist, and Dam.

HEMORRHAGIC DISEASE OF THE NEWBORN. Brinkhous, Smith, and Warner made a study of the prothrombin content of blood of the newborn and revealed the fact that, at birth, their value is considerably lower than that of the mother, and that the values decrease during the first few days of life but begin to rise again about one week after birth. Quick and Grossman believe that the delayed clotting during the first week of life could be due to time required to establish, in the previously sterile intestine, bacterial flora which can produce vitamin K. It has also been reported by Kato and Poncher that 30 to 50 per cent of deaths of the newborn are due to hemorrhage.

The high incidence of intracranial hemorrhage and bleeding at other sites, coupled with the discovery of low prothrombin values in the blood of the newborn, strongly indicates a therapeutic trial of vitamin K. It was shown conclusively that treatment of the mother during the terminal period of gestation, even as late as two hours before delivery, or of the infants immediately after birth, was effective in causing the production of prothrombin.

Several obstetricians have advocated the routine prophylactic use of vitamin K. However, the view that vitamin K is effective in preventing hemorrhage in the newborn has been challenged in five or six hundred papers on this subject. Sanford and co-workers studied over 1600 infants. The mothers of approximately one-half were given vitamin K antenatally. A comparison of this group with the controls revealed no difference in the incidence of hemorrhage. Sanford *et al.* showed that the blood of babies whose mothers had been treated with vitamin K had a higher prothrombin content than those of the untreated group.

Potter gave vitamin K to over 6000 women prior to delivery and compared the incidence of stillbirths and death of infants with an untreated group. There was very little difference in mortality. She remarks that "the incidence of intracranial hemorrhage is directly related to the skill and judgment of the obstetrician."

OTHER DISEASES. Experimentally, vitamin K seems to be useful only in facilitating the production of prothrombin. A beneficial therapeutic effect should, therefore, be expected in any disease in which there is a deficiency of vitamin K, if the patient has a liver with adequate function.

A number of physicians have found a lowered prothrombin in sprue, presumably due to impaired absorption accompanying the severe diarrhea. Administration of vitamin K caused a rise in the prothrombin content and cessation of bleeding.

As far as we know, vitamin K deficiency due to an inadequate intake is rare.

Both Day *et al.*, and Sebrell and associates have shown that prolonged administration of some of the sulfa drugs to rats causes severe hypoprothrombinemia, presumably due to the inhibitory effect on bacterial flora in the intestines.

The absorption of vitamin K has been prevented through production of a deficiency syndrome by giving mineral oil to rats daily. Javert has warned that daily ingestion of mineral oil by the human causes hypoprothrombinemia.

A number of investigators have proposed a new liver function test, based on the increase of the prothrombin of the blood following the administration of a fixed amount of vitamin K. Attempts to correlate other liver function methods with the new "prothrombin response test has given rise to conflicting results." The main function of the prothrombin test will remain as a guide to vitamin K therapy. It may yield valuable information concerning the pathological physiology of the liver. Allen has observed that the differential diagnosis between intrahepatic and obstructive jaundice can be made with a high degree of accuracy when one observes the plasma prothrombin response to a course of vitamin K therapy.

An interesting relationship between sweet clover disease and vitamin K has been discovered. Link, who has isolated the toxic principle, dicumarol, from spoiled sweet clover, has shown that this compound produces its effect by inhibiting the production of prothrombin, thereby leading to hypoprothrombinemia and fatal hemorrhage. Overman *et al.,* using experimental animals, showed that vitamin K will counteract the effect of dicumarol. Since dicumarol is used extensively in an attempt to prevent thrombosis, it is necessary to determine the prothrombin frequently. If it falls to a dangerous level, treatment with vitamin K must be instituted. Shapiro *et al.,* Lucia and Aggeler, and Cromer and Barker have shown that dicumarol can be used successfully in human patients.

Link and his associates have reported that the metabolism of dicumarol gives salicylic acid and, moreover, he has shown that administration of excessive amounts of salicylic acid produces hypoprothrombinemia. This should be kept in mind when massive salicylate therapy is employed, as in the Coburn treatment of rheumatic fever.

BIBLIOGRAPHY

Allen, J. G.
1943. Intern. Absts. Surg. **76**: 401
Almquist, H. J., & A. A. Klose
1939. Biochem. J. **33**: 1055.
1940. Proc. Soc. Exp. Biol. & Med. **45**: 55.
Almquist, H. J., C. F. Pentler, & E. Mecchi
1931. Proc. Soc. Exp. Biol. & Med. **38**: 336.
Almquist, H. J., & E. L. R. Stokstad
1936. J. Nutrition **12**: 329.
Ansbacher, S.
1939. J. Nutrition **17**: 303.
1940. Proc. Soc. Exp. Biol. & Med. **44**: 248.
Bancroft, F. W., I. N. Kugelmass, & M. Stanley-Brown
1929. Ann. Surg. **90**: 161.
Brinkhous, K. M., H. P. Smith, & E. D. Warner
1937. Am. J. Med. Sci. **193**: 475.
Butt, H. R., A. M. Snell, & A. E. Osterberg
1939. Proc. Staff Meet. Mayo Clinic **14**: 497.
Carr, J. L., & F. S. Foote
1934. Arch. Surg. **29**: 277.
Coburn, A. F.
1931. Williams & Wilkins Company. Baltimore.
Cromer, H. E., Jr., & N. W. Barker
1944. Proc. Staff Meet. Mayo Clinic **19**: 217.
Dam, H.
1929. Biochem. Z. **215**: 475.
1930. Biochem. Z. **220**: 158.
1935. Biochem. J. **29**: 1273.
Dam, H., & J. Glavind
1938. Lancet **234**: 720.
Dam, H., J. Glavind, & P. Karrer
1940. Helv. Chim. Acta **23**: 224.
Dam, H., & F. Schonheyder
1934. Biochem. J. **28**: 1355.
Dam, H., F. Schonheyder, & E. Tage-Hansen
1936. Biochem. J. **30**: 1075.

Dann, F. P.
1938. Am. J. Physiol. **123**: 48.
1939. Proc. Soc. Exp. Biol. & Med. **42**: 663.
Day, H. G., K. G. Wakim, M. M. Krider, & E. E. O'Banion
1943. J. Nutrition **26**: 585.
Elliot, M. C., B. Gaacs, & A. C. Ivy
1940. Proc. Soc. Exp. Biol. & Med. **43**: 240.
Foster, R. H. K., J. Lee, & U. V. Solmssen
1940. J. Am. Chem. Soc. **62**: 453.
Greaves, J. D., & C. L. A. Schmidt
1937. Proc. Soc. Exp. Biol. & Med. **37**: 43.
Greaves, J. D.,
1939. Am. J. Physiol. **125**: 429.
Halbrook, E. R.
1935. Thesis, University of California.
Holst, W. F., & E. R. Halbrook
1933. Science **77**: 354.
Javert, C. T., & S. Macri
1941. Am. J. Obst. & Gynec. **42**: 409.
Kato, K., & H. G. Poncher
1940. J. A. M. A. **114**: 749.
Kornberg, A., F. S. Daft, & W. H. Sebrell
1944. J. Biol. Chem. **155**: 193.
Link, K. P., R. S. Overman, W. R. Sullivan, C. F. Huebner, & L. D. Scheel
1943. J. Biol. Chem. **147**: 463.
Lucia, S. P., & P. M. Aggeler
1941. Am. J. Med. Sci. **201**: 326.
McFarlane, W. D., W. R. Graham, Jr., & F. Richardson
1931. Biochem. J. **25**: 358.
Morawitz, P., & R. Bierich
1907. Arch. Exp. Path. Pharmakol. **56**: 115.
Overman, R. S., M. A. Stahmann, & K. P. Link
1942. J. Biol. Chem. **145**: 155.
Potter, E. L.
1945. Am. J. Obst. & Gynec. **50**: 235.
Quick, A. J.
1935. J. Biol. Chem. **109**: lxxiii.
1937. Am. J. Physiol. **118**: 260.
Quick, A. J., & A. M. Grossman
1940. Am. J. Med. Sci. **199**: 1.
Ravdin, I. S., C. Riegel, & J. L. Morrison
1930. Ann. Surg. **91**: 801.
Sanford, H. N., I. Shmigelsky, & J. M. Chapin
1942. J. A. M. A. **118**: 697.
Schonheyder, F.
1935. Nature **135**: 653.
Sebrell, W. H.
1943-1944. Harvey Lectures: 288.
Shapiro, S., M. H. Redish, & H. A. Campbell
1943. Proc. Soc. Exp. Biol. & Med. **52**: 12.
Smith, H. P., & C. A. Owen
1940. J. Biol. Chem. **134**: 783.
Thayer, S. A., D. W. MacCorquodale, R. W. McKee, & E. A. Doisy
1938. J. Biol. Chem. **123**: CXX.
Thayer, S. A., R. W. McKee, S. B. Binkley, D. W. MacCorquodale, & E. A. Doisy
1939. Proc. Soc. Exp. Biol. & Med. **40**: 478.
Tidrick, R. T., F. T. Joyce, & H. P. Smith
1939. Proc. Soc. Exp. Biol. & Med. **42**: 853.
Warner, E. D., K. M. Brinkhous, & H. P. Smith
1938. Proc. Soc. Exp. Biol. & Med. **37**: 628.
Warner, E. D., & J. E. Flynn
1940. Proc. Soc. Exp. Biol. & Med. **44**: 607.
Whipple, G. H.
1913. Arch. Int. Med. **12**: 637.

HEMORRHAGIC MANIFESTATIONS OBSERVED IN EXPERIMENTAL DEFICIENCY OF PANTOTHENIC ACID, CHOLINE, AND CYSTINE

By PAUL GYÖRGY

Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania

Vitamin K, the specific "coagulation vitamin", is not the only nutritional factor related to hemorrhagic manifestations. The classical example of a nutritional deficiency disease, representing a characteristic hemorrhagic syndrome, is scurvy, caused by lack of vitamin C. More recently, the alleged participation of vitamin P and related substances in the maintenance of normal vascular permeability has been widely discussed. It would have been difficult to arouse any special interest for the generally well-known experimental and clinical picture of scurvy, and equally difficult to harmonize the rather contradictory experimental, clinical findings and various opinions concerning the role of so-called vitamin P in the prevention of hemorrhages of vascular origin. In the following, I shall confine myself to the review of experimental hemorrhagic conditions linked with members of the vitamin B complex, in which I have had a special personal interest for a long time.

It was about 10 years ago[1] when I first observed, in young rats kept on a synthetic diet, supplemented with the fat-soluble vitamins A and D as well as with thiamine, riboflavin, and pyridoxine, a very impressive syndrome of hemorrhagic diathesis. Rats, mostly albino, 21 days of age, weighing not more than 35 gm. were placed on the following diet: casein, 18; cane sugar, 68; melted butter fat, 9; salt mixture, 4; cod liver oil, 1.

Each rat utilized in an experiment received daily in addition to the above diet, supplements of thiamine (10—20 μg.), riboflavin (10—25 μg.), and pyridoxine (10—20 μg.). In many instances, the addition of pyridoxine was delayed until symptoms typical of acrodynia appeared. The hemorrhagic disease in question was observed in animals receiving only thiamine and riboflavin, but more often in animals supplemented with thiamine, riboflavin, and pyridoxine.

The hemorrhagic diathesis was indicated by variable signs. One of the most common symptoms was nosebleed. Melena and profuse hematuria were rather rare occurrences. The most impressive manifestations of the hemorrhagic diathesis in question were blood effusions into the skin exhibiting all the characteristics of purpura. In the gross, there were patches of purpura of varying size, some confluent, resulting in larger areas of dark red or blue discoloration. These occurred mainly on the dorsum of the feet and, particularly, on the hind legs or on the face and neck. As a rule, purpura is a late manifestation of this specific disease.

When it appears, it gradually increases in severity up to the fatal issue which normally follows in from 1 to 2 days.

Microscopically, in the regions of purpura, the epidermis varied in thickness. Extravasation of blood was present within the epidermis and in the corium immediately beneath the epidermis. There was no deposition of blood pigments, and little or no inflammatory reaction was apparent around the extravasated blood. The deeper portions of the corium and the subcutaneous tissue showed only occasional small foci of hemorrhage (FIGURE 1).

Examination of the blood of rats in the acute stage of hemorrhagic diathesis revealed, in the majority of the animals, low percentage of hemoglobin, and low platelet, white blood and red blood cell counts, with a very pronounced granulocytopenia. The clotting time was, as a rule, within normal limits.

From the foregoing summary of the hemorrhagic manifestations observed and the morphologic blood findings, it became evident that we were dealing with a profound disturbance of the primary blood-producing tissue, the reticuloendothelium, in its transformation into the three distinct types of blood cells, *i.e.,* into red blood cells, white blood cells, and megakaryocytes (platelets). Generally, the disease was ushered in by granulocytopenia, but was followed soon by thrombocytopenia, affecting the red blood cells only at a later stage in the form of progressive anemia. This combination of symptoms and their consecutive appearance are characteristic of the clinical entity called aleukia hemorrhagica (panmyelophthisis), which again is more or less synonymous with the term aplastic anemia.

Fresh biopsy specimens of bone marrow (femur, humerus, vertebrae) from healthy animals showed that marrow was very cellular and included cells of all three blood systems: white blood cells, red blood cells, and megakaryocytes. Appearance of the marrow in animals which succumbed to the experimental aleukia hemorrhagica varied in different bones and in the bones of different animals. In some, the only abnormalities were intense hyperemia, almost entire absence of megakaryocytes, and a reduction in the number of cells, chiefly the granulocytes, in the marrow. In others, in addition to the severe hyperemia, there was a varying amount of hemorrhage and edema, accompanied by more pronounced reduction in the number of cells in the marrow, and the granulocytes and megakaryocytes were altogether missing. Many of the patches of edema contained deposits of fibrin (FIGURE 2). In some animals, the hematopioetic cells in the tissue of the marrow had almost completely disappeared.

Necrotic and ulcerative lesions around the month and on the mucous membranes, almost specific signs of human agranulocytosis, were observed rather infrequently in rats with granulocytopenia.

A frequent *post-mortem* finding in the internal organs, and the most striking, was hvperemia and hemorrhage in the suprarenal bodies. The hemorrhage, when present, was easily recognizable and varied in extent from small patches in the cortex or medulla, or both, to diffuse hemorrhage with obliteration of the natural architecture of the entire organ.

Microscopically, in the regions of hyperemia alone, the vessels were distended with blood and the cells between them showed varying degrees and amounts of degenerative change. In those parts where the hyperemia was intense and, also, in the regions of extravasation of blood, the parenchyma was almost or entirely destroyed (FIGURE 3). Necrosis may also be found without concomitant hemorrhage.[1-4]

In several animals that had clinical melena, the large intestine showed, microscopically, some extravasation of blood in the mucosa without accompanying inflammation.

At macroscopical examination, the spleen showed only hyperemia. Microscopically, the sinusoids were distended with blood. In the pulp, especially around the lymphoid follicles, there were zones of hemorrhage. Most of the follicles had no germinal centers. The most striking abnormality was the absence of megakaryocytes. In the pulp, there was moderate cellularity with a varying amount of histiocytic hyperplasia and fibrosis.

In the search for the probable cause of the hemorrhagic syndrome described above, we were unable to relate it to an infectious agent. The fact that it was observed among rats kept on a special deficient diet and that it could be prevented by supplements of yeast, yeast extract, liver extract, milk, etc., strongly supported the assumption of nutritional etiology.

In animals in which aleukia hemorrhagica was diagnosed at an early stage, administration of a yeast extract preparation prevented the fatal issue and restored normal conditions. The animals showed gradual improvement in the morphologic blood composition and in the permeability of the vessels (purpura). Unfortunately, many animals were already in a moribund state when the very acute, usually almost fulminant breakdown in hematopoiesis was discovered. They died despite treatment, from 1 to 2 days after it had been initiated. In the successfully treated animals the histologic picture of the bone marrow (FIGURE 4) and spleen became normal, showing the usual large number of megakaryocytes.

At the time of our original observations, no pure pantothenic acid was available. In view of the fact that an alleged crude concentrate of pantothenic acid (at that time called "filtrate factor") was found to be ineffective, we first believed that this factor could not be the determining cause of the specific hemorrhagic conditions.

In the course of our continued investigations, all of which remained so far unpublished, the following additional facts were established. (1) In a group of 250 rats kept on the routine B-free diet supplemented with thiamine, riboflavin, and pyridoxine, 46 rats died with signs of massive adrenal hemorrhage without concomitant involvement of the hematopoietic system, 28 rats died from aleukia hemorrhagica without hemorrhagic manifestations in the suprarenal bodies, and 13 rats showed both aleukia hemorrhagica and adrenal hemorrhages. (2) In a group of 62 rats receiving the same basal diet, supplemented first with thiamine and riboflavin, and later, after the appearance of signs of rat acrodynia, with pyridoxine and pure synthetic calcium pantothenate, 49 rats remained free from hemorrhagic manifestations, 2 animals died from adrenal hemorrhages 6 and 14 days after the start of the pantothenic acid medication, one animal died with purpura and adrenal hemorrhage 15 days, and 10 rats died from aleukia hemorrhagica between 5 and 18 days (on the average, 10½ days) after pantothenic acid was first given. (3) In a group of young weanling rats, comprising more than 300 animals receiving the same basal diet supplemented with thiamine, riboflavin, pyridoxine and calcium pantothenate from the start, not one animal developed either aleukia hemorrhagica or adrenal hemorrhages.

All these observations seem to support the view that both adrenal hemorrhages and aleukia hemorrhagica are related to pantothenic acid. On the other hand, their relatively low incidence even under severe conditions of deficiency of pantothenic acid, and their occurrence up to 18 days after pantothenic acid was supplied to animals already in the latent stage of the disease, appear to be in accord with the conclusion that pantothenic acid acts not directly but through the intermediary of some other vitamin or metabolite.

Since our first observations, the presence of adrenal hemorrhages in animals kept on a synthetic diet free from pantothenic acid and its prevention by pantothenic acid has been confirmed in several other laboratories.[5-9] It has also been shown, by other workers, that anemia and granulocytopenia may be encountered and prevented in rats under similar dietary conditions.[10-12] Folic acid proved to be ineffective in the prevention of the specific blood dyscrasia in rats kept on a diet deficient in pantothenic acid.[11]

There is a major discrepancy between our own observations and those made in several other laboratories, *i.e.*, the relatively high incidence of purpura and thrombocytopenia in our animals and their absence in rats kept under similar dietary conditions in several other laboratories. It is conceivable that this difference may be due to variations in quantity and also quality of fat in the experimental rations used.

Purpura, suprarenal hemorrhage, prostration, and a rapidly fatal course seen in our rats, also characterize the clinical picture of the so-called Waterhouse-Friderichsen (W-F) syndrome. In this connection, it is of special interest that in a case of the W-F syndrome, Glanzmann found progressive thrombocytopenia and distinct diminution of the granulocytes with degenerative changes in the remaining polymorphonuclear cells.[13] Although these blood findings seem to be rather exceptional, and leucocytosis prevailed in similar cases, the close analogy between the W-F syndrome and the appearance of some of our rats that were ill with panmyelophthisis can, nevertheless, be regarded as exceedingly striking.

Etiology of the W-F syndrome is now generally attributed to a septic infection and, in particular, certainly in the majority of cases, to a fulminant meningococcus sepsis which ends fatally before the local alterations of meningitis may develop. However, in rats ill with the corresponding syndrome, we were unable to prove the existence of a primary infection as a decisive etiologic factor. Whereas a negative observation such as this does not exclude, with certainty, the presence of an unrecognized infection, the nutritional control of the production of the syndrome in rats permits us to make at least one other suggestion with regard to its human equivalent. From the rat experiments, we may hypothesize that the bacterial, septic etiology of the W-F syndrome in man is built up on the basis of a nutritional deficiency similar to that which was found necessary for the production of the analogous disturbance in rats. With this assumption, it becomes conceivable why meningococci provoke the specific W-F syndrome only in certain persons.

One could also look upon adrenal hemorrhage and adrenal necrosis as the "stage of exhaustion" during the course of an alarm reaction. Its presence in rats kept under special dietary conditions might conceivably indicate a close relationship between pantothenic acid and functional state of the adrenal cortex. Such possibility has been discussed from a somewhat different angle in the past[4, 14] and requires further intensive study.

The hemorrhagic syndrome in experimental deficiency of pantothenic acid in rats is due mainly to an underlying general blood dyscrasia. In contrast, hemorrhagic manifestations occurring in experimental deficiency of choline and in related conditions are caused probably by direct damage of the organs involved, especially the kidney and liver. Here, disturbed synthesis and metabolism of phospholipids may be the leading pathogenic factor.

Fat infiltration of the liver as a regular manifestation of lack of choline has been known for a long time. More recently, Griffith and his collaborators,[15, 16] as well as work from our laboratory,[17] in addition to various other later studies of several other workers, have established a direct relationship of severe, often fatal hemorrhagic and necrotic lesions in the kidneys in rats with absolute or relative deficiency of choline (relative to the cystine/methionine quotient of the diet).

The so-called "hemorrhagic kidney" is an almost regular occurrence in young weanling rats fed a diet poor in methionine and free from choline. Rats kept on such basal ration containing a sufficient amount of protein and of adequate caloric value usually succumb to the renal changes between the 11th and 15th day of the experiment. In the acute stage of the disease, the kidneys are large, with a purplish-red surface color, apparently due to intracapsular and subcapsular hemorrhage. On section, the cortex is diffusely dark red, or there are alternating patches of dark red and yellow throughout the cortex, in contrast to the grayish-yellow color of the medulla. The color of the surface and of the cross sections of the liver is, as a rule, uniformly light yellow, suggesting a high content of fat.

Microscopically, between the fibers of the connective tissue in the capsule and also beneath the capsule, there is a variable amount of recently extravasated blood. The most pronounced lesions of the parenchyma consist of almost complete necrosis of the tubules of the cortex (acute necrotizing nephrosis), without destruction of the architecture. In most kidneys, the outline of the tubules is still recognizable. A striking feature of the diffuse necrosis of the cortex is the intensive hyperemia of the interstitial tissue between the tubules and surrounding the glomeruli (FIGURE 5). The hyperemia is so intense in the cortex of some of the kidneys that it has the appearance of hemorrhage and imparts the dark red color to the tissue observed in the gross, merging with the hemorrhagic layer in the subcapsular space.[17-19]

Choline deficiency, with a relatively high cystine/methionine quotient in the diet, will aggravate the hemorrhagic changes in the kidneys. In rats receiving a diet low in choline (methionine) and high in cystine, hemorrhagic manifestations may also be observed in the eyes, heart muscle, liver (FIGURE 6), and in the adrenals.[19-22] Not only a relatively high[23-27] but also a relatively very low ratio of cystine/methionine and choline is conducive to the development of patchy hemorrhagic necrosis in the liver.[28]

"Hemorrhagic kidney" may also be encountered, although only as an irregular finding, in older rats kept on a synthetic ration low in choline and relatively high in cystine. In general, however, necrotizing nephrosis without concomitant hemorrhagic manifestations is the rule in this older group of animals.

It is of especial interest that hemorrhagic necrosis in the liver as well as "hemorrhagic kidney" are found, identical experimental conditions notwithstanding, more often in male than in female animals. The protective effect of estrogenic hormones in the liver and kidney may also be demonstrated in various other pathological reactions.[27]

The hemorrhagic and necrotic lesions in the liver of rats bear some resemblance to the acute and subacute stages of yellow liver atrophy in man. Bilateral hemorrhagic cortical necrosis of the kidneys seen in rats is known as a clinical entity. Thus, both conditions may be identical or at least closely related to each other. One possible difference between the microscopic pictures of the two concerns vascular thrombosis, which has been described in the human kidneys but was not observed in the kidneys of the animals. The pathogenesis of the renal lesions in the animals must, therefore, be attributed to a nephrotoxic effect acting directly on the parenchyma, and not to ischemia brought about by vascular thrombosis. Since thrombosis is not always found even in the human kidneys, there remains the possibility that bilateral symmetrical hemorrhagic cortical necrosis in human beings is also due to a nephrotoxic effect, with or without angiospasm, determined by a dietary deficiency, and that the thrombosis is only a terminal event that plays no primary part in the pathogenesis of the renal lesions.[17]

The hemorrhagic lesions in the liver and kidneys are the fulminant form of the so-called "hepato-renal syndrome." Its possible dietary origin and control may have practical significance beyond the specific pathologic condition in question, and may extend to the humoral regulation of blood pressure, including shock prevention and vascular reactions to hemorrhage.

BIBLIOGRAPHY

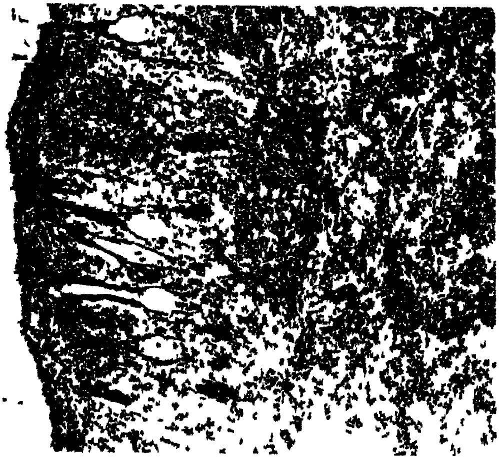
1. **György, P., H. Goldblatt, F. R. Miller, & R. P. Fulton**
1937. J. Exp. Med. **66**: 579.
2. **Nelson, A. A.**
1939. Pub. Health Rep. **54**: 2250.
3. **Ashburn, L. L.**
1940. Pub. Health Rep. **55**: 1337.
4. **Deane, H. W., & J. M. McKibbin**
1946. Endocrinology **38**: 385.
5. **Daft, F. S., & W. H. Sebrell**
1939. Pub. Health Rep. **54**: 2247.
6. **Daft, F. S., W. H. Sebrell, S. H. Babcock, Jr., & T. H. Jukes**
1940. Pub. Health Rep. **55**: 1333.
7. **Unna, K., & W. L. Sampson**
1940. Proc. Soc. Exp. Biol. & Med. **45**: 309.
8. **Mills, R. C., J. H. Shaw, C. A. Elvehjem, & P. H. Phillips**
1940. Proc. Soc. Exp. Biol. & Med. **45**: 482
9. **Salmon, W. D., & R. W. Engel**
1940. Proc. Soc. Exp. Biol. & Med. **45**: 621.
10. **Daft, F. S., A. Kornberg, L. L. Ashburn, & W. H. Sebrell**
1945. Pub. Health Rep. **60**: 1201.
11. **Daft, F. S., & W. H. Sebrell**
1946. Federation Proc. **5**: 231.
12. **Carter, C. W., R. G. Macfarlane, J. R. P. O'Brien, & A. H. T. Robb-Smith**
1945. Biochem. J. **39**: 339.
13. **Glanzman, E.**
1933. Jahrb. f. Kinderh. **139**: 49.
14. **Selye, H.**
1946. J. Clin. Endocrinol. **6**: 117.
15. **Griffith, W. H. & N. J. Wade**
1939. J. Biol. Chem. **131**: 567.
16. **Griffith, W. H.**
1941. Biological Symposia **5**: 193.
17. **György, P., & H. Goldblatt**
1940. J. Exp. Med. **72**: 1.
18. **Christensen, K.**
1940. J. Biol. Chem. **133**: XX.
19. **Christensen, K.**
1942. Arch. Path. **34**: 630.
20. **Engel, R. W., & W. D. Salmon**
1941. J. Nutrition **22**: 109.
21. **Jervis, G. A.**
1942. Proc. Soc. Exp. Biol. & Med. **51**: 193.
22. **Bellows, J. G., & H. Chinn**
1943. Arch. Ophth. **30**: 105.
23. **Curtis, A. C., & L. H. Newburgh**
1927. Arch. Int. Med. **39**: 829.
24. **Sullivan, M. X., W. C. Hess, & W. H. Sebrell**
1932. Pub. Health Rep. **47**: 75.
25. **Lillie, R. D.**
1932. Pub. Health Rep. **47**: 83.
26. **Earle, D. P., Jr., & J. Victor**
1941. J. Exp. Med. **73**: 161.
27. **György, P., & H. Goldblatt**
1942. J. Exp. Med. **75**: 355.
28. **Weichselbaum, T. E.**
1935. Quart. J. Exp. Physiol. **25**: 363.

PLATE 1

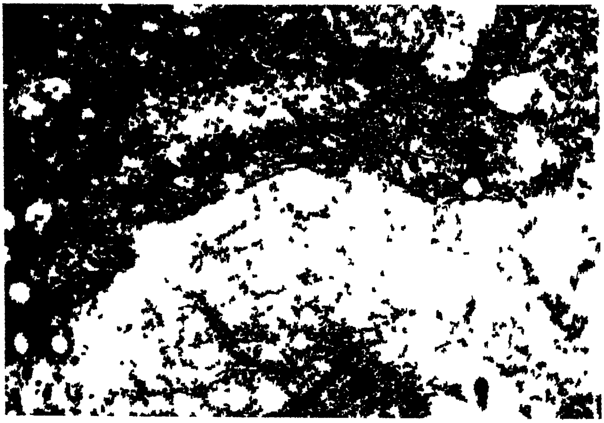
FIGURE 1. Section of skin, showing hyperemia, hemorrhage, and edema in upper portion of corium. x 100

FIGURE 2. Section of bone marrow. Varying amount of hemorrhage and edema is shown, with deposits of fibrin in many of the patches of edema. x 297.

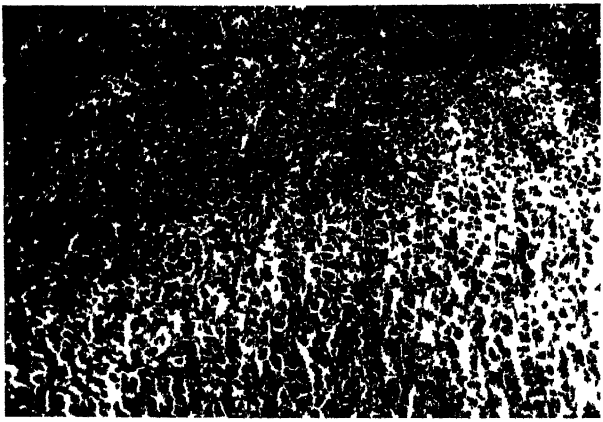
FIGURE 3. Section of cortex of suprarenal body, showing degeneration of cortical cells in the region of hemorrhage and intense hyperemia. x 199.



1

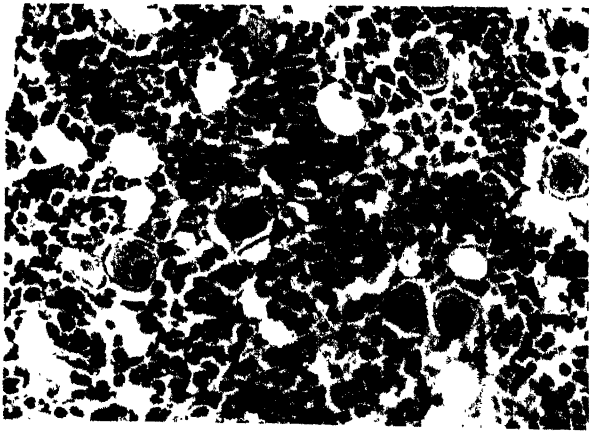


2

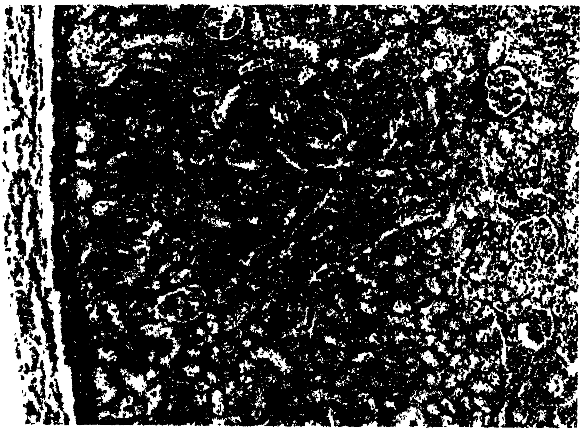


3

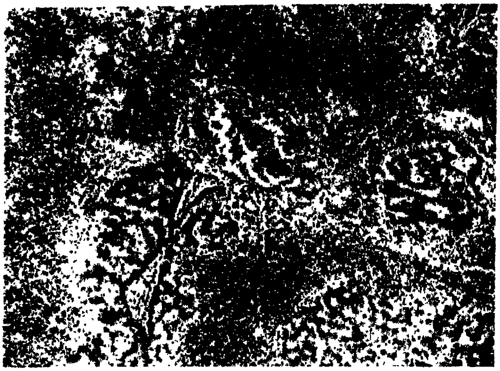
GYORGY: HEMORRHAGE IN PANTOTHENIC ACID DEFICIENCY



4



5



6

GYORGY: HEMORRHAGE IN PANTOTHENIC ACID DEFICIENCY

PLATE 2

FIGURE 4. Section of bone marrow after treatment and complete recovery. There is a normal number of megakaryocytes, and the cellular marrow shows many granulocytes. x 570.

FIGURE 5. Kidney. Acute diffuse necrosis of tubules of cortex and intense hyperemia of the interstitial tissue. There is hemorrhage into and beneath the renal capsule and some interstitial extravasation of blood in the periphery of the cortex. x 150.

FIGURE 6. Extensive necrosis of liver cells and much extravasated blood. Only a few liver cells still recognizable around a portal space. x 87.

THE EFFECTS OF HEMORRHAGE ON THE CIRCULATION*

By Dickinson W. Richards, Jr.

Department of Medicine, Columbia University, College of Physicians and Surgeons, and the Chest Service, Bellevue Hospital, New York, N. Y.

In the first World War, after prolonged investigation, the emphasis in traumatic and hemorrhagic shock was finally placed upon mechanical causes, particularly loss of blood volume. Starting with this point of view in World War II, we have found the basic mechanical facts, though they have been confirmed, inadequate to account for the ultimate tissue disturbances, and have therefore come to seek more fundamental explanations in vasomotor and chemical phenomena. We are, in this respect, returning to the point of view of fifty years ago.

With this brief introduction, I should like to present a description of recent observations in clinical shock due to hemorrhage, leading up to the unknown factors, especially in its so-called irreversible stage.

Blalock,[1] in his excellent review written in 1940, concisely stated the case for the mechanical concept of shock, about as follows: The most important alteration in traumatic shock is decreased blood volume, the consequences of this being diminished venous return and decreased cardiac output. Vasoconstriction at first holds up blood pressure, but with further blood loss this also declines. Work by Blalock,[2] Phemister,[3] and others had demonstrated the important fact that in trauma produced in various ways, the decrease in blood volume is largely or entirely due to fluid loss at the site of injury.

Taking up the problem at this point, in a series of clinical studies carried out in the early years of the war, with the use of cardiac catheterization and other techniques, it was possible for several groups of investigators to confirm this general concept[4, 5, 6] and also to extend it, by defining separately the state of the circulation in various different forms of injury with shock.

Of these, the simplest, *i.e.*, the form least affected by complicating factors, was external hemorrhage, a large loss of blood taking place progressively. Figure 1 shows the basic findings, as compared with the normal state and with a series of cases of shock due to skeletal trauma, in twelve cases of hemorrhage. It will be seen that blood volume is reduced by about two-fifths, with red cell loss proportionately greater than that of plasma, right auricular pressure or filling pressure of auricle and ventricle much below normal, and cardiac output about two-thirds to one-half normal. Arterial blood pressure and pulse pressure are decreased; peripheral resistance, which is the ratio of mean arterial pressure to cardiac output, is normal or slightly increased.

---

* The work reported in this paper was supported, in part, by a grant from the Commonwealth Fund.

The importance of hemorrhage in traumatic shock was enhanced when clinical findings were shown to support those of animal work in indicating that in other traumatic injuries, such as multiple fractures, the deficit in blood volume was actually a loss of whole blood (= hemorrhage) at the site of injury. Correspondingly, the mechanism of shock in trauma is almost identical with that of hemorrhage. This also is shown in FIGURE 1.

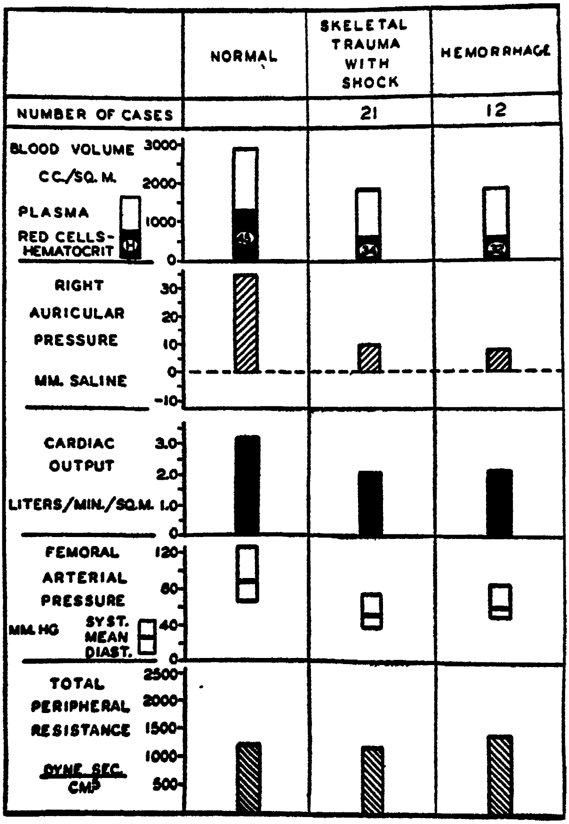


FIGURE 1. Chart showing chief dynamic changes in the circulation in traumatic and hemorrhagic shock. Average measurements.

As a corollary to this general picture, we have the fact, now well known, that in both traumatic and hemorrhagic shock, there is hemodilution, with low hematocrit. Gregersen and Noble[7] were able to show, by successive blood volume measurements, that this relative increase in plasma was a compensatory phenomenon, fluid and protein coming into the blood from uninjured tissues, following after the gross loss of whole blood at the site of injury.

From these findings also, it was more than obvious that, for treatment of shock, the primary need was whole blood rather than plasma. This was established as early as the spring of 1942.

Carrying the consequence of shock one step further, to the tissues themselves, it seemed at first that this mechanical concept was adequate to explain the ultimate tissue anoxia which constituted the final breakdown in the vital processes. Certainly, with decreased total blood flow, and decreased oxygen-carrying capacity of each unit of blood, due to hemodilution, the total oxygen transport to the tissues is much diminished. More careful study of the quantitative changes, however, and especially a comparison with other forms of shock, showed that the tissue anoxia could not very well be explained so simply, and that other disturbances in the peripheral vascular system must be considered. I shall return to this later.

The first comprehensive study, however, of actual vasomotor failure, was that carried out by McMichael,[8] and later also by Stead,[9] in a series of cases studied during and following a rapid phlebotomy, during which the subject fainted. The clinical picture of syncope was, of course, well known, as produced by tilt table and other procedures; but the measurement of the whole circulatory dynamics was new. McMichael found, as the removal of blood proceeded, a small decrease in cardiac output and right auricular pressure. When the subject fainted, the sudden and profound drop in arterial pressure was not associated with any further drop in cardiac output or in auricular pressure. McMichael also measured blood flow in the arm during the faint, and found this increased. Unmistakably, the conclusion was that the essential failure in syncope is sudden loss of vasomotor tone, especially in the blood vessels of skeletal muscles, without necessarily any significant drop in total blood flow, *i.e.*, cardiac output. Further proof was provided by McMichael through the administration of methedrine during the faint, with immediate recovery of the circulation to normal.

Relative magnitude of the changes observed in this experiment will vary with the particular technique, especially speed of removal of blood and amount removed. FIGURE 2 shows a record of pressure tracings from the right ventricle (lower tracing) and femoral artery (upper tracing) before, during, and after a large and rapid phlebotomy. At the time of fainting, in this instance, pressures were much diminished, as shown. However, cardiac output was also decreased, being only 40 per cent of its normal level. The subsequent records show restoration of the circulation after an infusion.

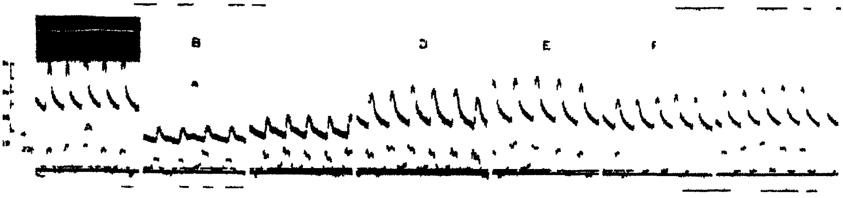
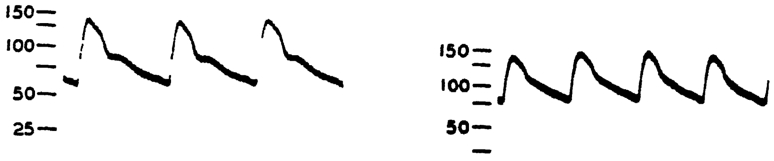


FIGURE 2. Tracings from femoral artery and from right ventricle in a normal subject before (A) and immediately after (B) a large phlebotomy. Patient was pale and in shock. Cardiac output was 40 per cent of normal. C, D E, and F = progressive stages in recovery during and following an intravenous infusion of gelatin solution. G = 4 hours after phlebotomy. Cardiac output then 95 per cent of normal.

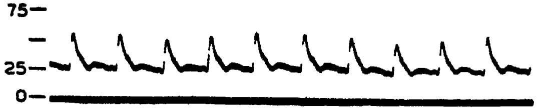
It should be noted that, in these studies, one starts with a normal individual with well-filled vascular bed, and particularly with well-filled great vessels, both venous and arterial. Stead[9] has made some interesting studies in this connection, showing that normal subjects can vary considerably their right auricular pressures with little or no change in cardiac output. In other words, a well filled venous reservoir provides an excellent reserve to insure adequate return of blood flow to the heart. At the other end of the spectrum is the patient with distended or overfilled vascular bed, as in congestive heart failure. In such cases, phlebotomy[10] or the application of venous cuffs,[11] by relieving the distended and dilated heart, may result in an increase in cardiac output.

Let us return to a further consideration of the empty vascular bed and diminished circulation of secondary hemorrhagic shock. One of the findings in these clinical studies has been that there is *selective* vasoconstriction, a compensatory phenomenon. Blood flow through the extremities, for example, and through the kidneys,[12] is diminished out of proportion to the decrease in total blood flow, thus saving blood for more immediately vital regions.

The relatively empty vascular system, held together, so to speak, by powerful constrictor action, is in a peculiarly unstable and precarious state. FIGURE 3 shows pressure tracings from femoral artery in a case of shock compared with two normal subjects. The collapsing quality of the pulse wave is well demonstrated. Clinically, it is well known that, in such patients, it takes very little to send them into serious or abruptly fatal collapse. This most commonly occurs by sudden loss of vasomotor tone. In our series of cases at Bellevue Hospital, we observed many such instances, two of them fatal. One was a young man with acute alcoholism, lacerated scalp, and considerable loss of blood. He was in moderate shock. He sat up in bed, lost consciousness almost immediately, and



I.N. - NORMAL F.M. - NORMAL



H.S. - SEVERE TRAUMATIC SHOCK

FIGURE 8 Femoral artery pressure tracings in normal subjects and in shock

expired within a few minutes. The second case was a man in severe shock, having a ruptured spleen, fractured pelvis, and dislocated ankle. An attendant manipulated his ankle very slightly, whereupon the patient broke out into a sweat, became pulseless, lost consciousness, and died fifteen minutes later. In many similar episodes, patients were saved by immediate rapid transfusion or infusion.

Granting, then, the extreme precariousness of the circulation in this advanced stage of secondary hemorrhagic and traumatic shock, we have, among other things, two important questions to answer: (1) Do the alterations in circulatory dynamics, as we have described them earlier in this discussion, in cases of sustained shock of moderate degree, provide adequate explanation for all the progressive changes that follow as shock becomes deeper? (2) Does adequate treatment by blood replacement always restore the circulation, or, if it does not, what are the circumstances of ultimate failure?

As to the first, we find cardiac output in moderate hemorrhagic shock about two-thirds to one-half normal. That is, in itself, by no means lethal, not even particularly harmful. Many cardiacs exist for years, in partial ambulatory activity, with cardiac output less than this. We find hematocrit decreased in acute shock, but again not more than is tolerated in many cases of secondary anemia. Blood pressure, too, is low but not lethally decreased in early and middle stages of shock. If we put all these factors together, there will be, it is true, a badly damaged circulation. The combination is that of an empty vascular system, with decreased hemoglobin concentration per unit of blood flowing, decreased total flow, at decreased pressure. Total oxygen brought to all the tissues is thus about one-quarter normal. By compensatory vasoconstriction in non-vital regions, as already described, oxygen transport to vital regions is raised, perhaps to half normal or thereabouts. The capillaries themselves are probably inadequately filled, per unit of tissue mass. In this connection, one common misconception might be mentioned. It is often stated that, since all the capillaries in the body, under normal conditions, hold only about 60 cc. of blood, even extreme blood loss need not seriously empty them. Such a concept neglects the obvious fact that in order to serve and maintain 1 cc. of blood in active capillaries, there is required an entire arborization of filled vessels, many times greater in volume, on both arterial and venous sides.

Summarizing this argument, we may suppose that in secondary shock due to hemorrhage or skeletal trauma, there is enough mechanical deficiency in the circulation to account for the progressing tissue anoxia that occurs. Other forms of capillary dysfunction are, of course, not excluded.

Consider for a moment, however, two other types of injury with shock, namely, abdominal injuries and burns. FIGURE 4 shows the circulation in burns; and that in abdominal injuries is similar. Here cardiac output is, to be sure, decreased, but there is hemoconcentration, and arterial oxygen transport to the tissues is thus not seriously diminished. And yet, in these injuries just as much as in hemorrhage, the tissues themselves give evidence, at this stage, of progressing anoxia. One finds, for example, rapid increase in lactic acid and decreasing arterial and venous serum pH. It would seem that, in addition to the disturbed general blood flow, there are probably serious defects or disorganization in tissue capillary function. If one were to try to reach beyond the overall figures as we find them for pressure and blood flow in secondary traumatic shock, one might say that among the essential features in the early and middle stages are (a) empty great vessels, and (b) disorganized capillary function, due to lack of filling or other disturbance. That is about as far as one can go clinically. The very interesting findings in animal experimentation, describing both capillary and tissue changes, especially those of liver and kidney, by Fine,[13] Long,[14] Knisely,[15] Zweifach,[16] Shorr,[17] and others,

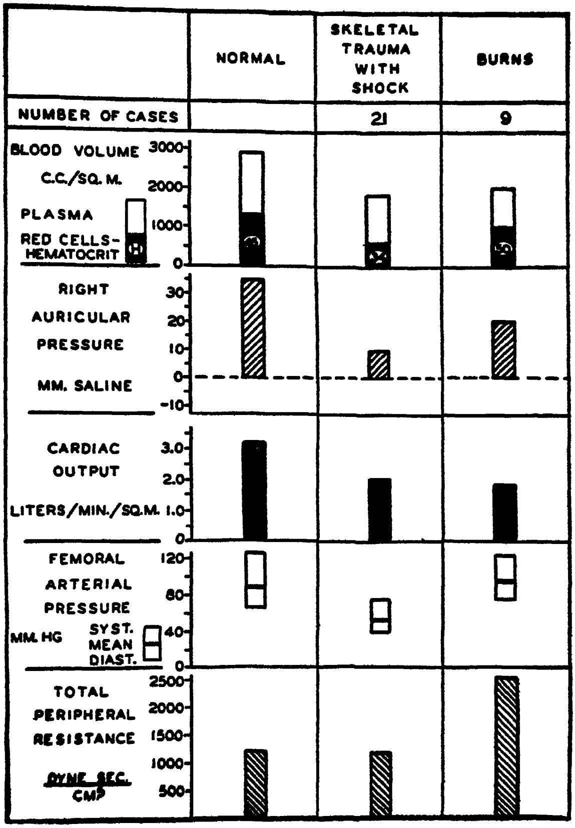


FIGURE 4. Chart showing chief hemodynamic changes in shock due to burns. Average measurements.

are probably applicable to the human problem, and may well give the explanation we want. This is, however, outside the scope of the present brief discussion.

Finally, how good is treatment by blood replacement? It needs no emphasis, at this time, to record the extraordinary recovery that can take place, even in advanced shock, by abundant replacement with plasma or whole blood. Cardiac output, right ventricular and arterial pressures all return to or near normal, and within an hour or two, blood lactic acid is well down and serum pH within normal limits. It is also, of course, inevitable that in some cases, injury is too severe, or shock has been too prolonged, for the patient to respond to treatment. In still others, there will be a temporary resuscitation, followed by a gradual failure again and death.

The nature of the ultimate failure is variable. If blood is poured in rapidly, there is almost always some initial response by the heart, and cardiac output and arterial pressure pick up. In a few cases in our series at Bellevue Hospital, respiratory failure occurred suddenly, while the circulation seemed to be improving. These were apparently cerebral deaths. Other forms of failure, occurring on the second or third day or even later, included pulmonary, renal, and occasionally hepatic manifestations, but these need not be considered here. More common, especially in severe battle injuries with prolonged shock, was the picture analogous to irreversible shock in animals, an early favorable response to transfusion, but followed by progressively fading pulse and blood pressure, in spite of sufficient blood to keep the blood volume up to normal. In this situation, as has been found abundantly in animal experimentation, no form of therapy has been effective. Wiggers[18] has raised the question whether, at this time, the ultimate failure may be cardiac rather than peripheral. One finds no way of answering this question from clinical evidence. There is nothing to indicate cardiac failure of the congestive type, though this does not rule out myocardial inadequacy. The general impression is that one is dealing, at this stage, with an inert vasomotor system and dying tissues,[19, 20] which may be myocardial as well as other. It is difficult to analyze the clinical problem further than this, with present evidence. While one may hope that something may be found to restore the vascular system and the tissues even at this stage, it will have to be a very powerful agent indeed.

By way of summary, it might be stated that, although recent clinical studies of secondary hemorrhagic shock have confirmed the basic mechanical alterations in blood flow found previously in animals, for adequate explanation of the whole picture the vasomotor aspects must be given emphasis. Further, in the so-called irreversible stage, the evidence suggests an almost total failure of response of both blood vessels and tissues. Shock is, in fact, very much more than disturbance in blood flow. Writing in 1872, S. D. Gross[21] defined shock as "the manifestation of a rude unhinging of the machinery of life." While lacking in detail, I am inclined to believe that Dr. Gross had the main idea.

BIBLIOGRAPHY

1. **Blalock, Alfred**
   1940. Principles of Surgical Care, Shock, and Other Problems. C. V. Mosby Co. St. Louis.
2. **Blalock, Alfred**
   1930. Experimental shock. The cause of the low blood pressure produced by muscle injury. Arch. Surg. **20:** 959.
3. **Parsons, E., & D. B. Phemister**
   1930. Hemorrhage and "shock" in traumatized limbs. Surg. Gyn. & Obst. **51:** 196.
4. **Cournand, A., R. L. Riley, S. E. Bradley, E. S. Breed, R. P. Noble, H. D. Lauson, M. I. Gregersen, & D. W. Richards, Jr.**
   1943. Studies of the circulation in clinical shock. Surgery **13:** 964.
5. **Evans, E. I., M. J. Hoover, G. W. James, & Theodore Alm**
   1944. Studies on traumatic shock. I. Blood volume changes in traumatic shock. Ann. Surg. **119:** 64.
6. **McMichael, John**
   1944. Clinical aspects of shock. J. A. M. A. **124:** 275.
7. **Gregersen, M. I., & R. P. Noble**
   1946. Blood volume in clinical shock. II. The extent and cause of blood volume reduction in traumatic, hemorrhagic, and burn shock. J. Clin Invest. **25:** 172.
8. **Barcroft, H., O. G. Edholm, J. McMichael, & E. P. Sharpey-Schafer**
   1944. Posthaemorrhagic fainting. Study by cardiac output and forearm flow. Lancet **1:** 489.
9. **Warren, J. V., E. S. Brannon, E. A. Stead, Jr., & A. J. Merrill**
   1945. The effect of venesection and the pooling of blood in the extremities on the atrial pressure and cardiac output in normal subjects. J. Clin. Invest. **24:** 337.
10. **Cournand, A.**
    Unpublished observations.
11. **McMichael, J., & E. P. Sharpey-Schafer**
    1944. The action of intravenous digoxin in man. Quart. J. Med. **13:** 123
12. **Lauson, H. D., S. E. Bradley, & A. Cournand**
    1944. The renal circulation in shock. J. Clin. Invest. **23:** 381.
13. **Frank, H. A., A. M. Seligman, & J. Fine**
    1946. Traumatic shock. XIII. The prevention of irreversibility in hemorrhagic shock by viviperfusion of the liver. J. Clin. Invest. **25:** 22.
14. **Wilhelmi, A. E., J. A. Russell, F. L. Engel, & C. N. H. Long**
    1945. Some aspects of nitrogen metabolism of liver tissue from rats in hemorrhagic shock. Am. J. Physiol. **144:** 674.
15. **Knisely, M. H., T. S. Eliot, & E. H. Bloch**
    1945. Sludged blood in traumatic shock. I. Microscopic observations of the precipitation and agglutination of blood flowing through vessels in crushed tissues. Arch. Surg. **51:** 220.
16. **Zweifach, B. W., R. G. Abell, R. Chambers, & G. H. A. Clowes**
    1945. Role of decompensatory reactions of peripheral blood vessels in tourniquet shock. Surg. Gyn. & Obst. **80:** 593.
17. **Shorr, E., B. W. Zweifach, & R. F. Furchgott**
    1945. On occurrence, sites, and modes of origin and destruction, of principles affecting compensatory vascular mechanisms in experimental shock. Science **102:** 489.
18. **Wiggers, C. J.**
    1942. The present status of the shock problem. Physiol. Rev. **22:** 74.
19. **Warren, J. V., & E. A. Stead, Jr.**
    1945. Orientation to the mechanisms of clinical shock. Arch. Surg. **50:** 1.
20. **McMichael, J.**
    1945. Present status of the clinical problem of "shock." Brit. Med. Bull. **3:** 105.
21. **Gross, S. D.**
    1872. System of Surgery. H. C. Leas Son & Co. Philadelphia.

EXPERIMENTAL STUDIES ON TRAUMATIC AND HEMORRHAGIC SHOCK

By Magnus I. Gregersen
Department of Physiology, College of Physicians and Surgeons,
Columbia University, New York, N. Y.

When one attempts to investigate a clinical problem such as shock through experiments in the laboratory, the first task is, or should be, to seek adequate proof that the clinical condition is actually reproduced in the experimental animal. The fact that we considered this rather seriously at the start of our investigations[1] of traumatic shock proved to be extremely fortunate. In order to reproduce in dogs the clinical picture of shock as described in man, we had to exclude the effects of general anesthesia, and by so doing we at once spared ourselves a difficulty which would have become increasingly oppressive as the investigations progressed and might well have blocked the exploration in several important directions, namely, the difficulty of differentiating between the effects of anesthesia and those produced by injury. Furthermore, the attention which we gave to the clinical signs[1]—if we may be permitted to apply this bedside terminology to dogs—made them valuable as indications of the differences in the pattern of response to different types of injury. The data that we gathered were constantly oriented with reference to the ordinary clinical criteria of shock, which enabled us to establish a time relation between certain bodily disturbances and the appearance of the clinical signs of shock.

It is hardly possible to overemphasize the importance of excluding the complicating effects of general anesthesia from studies on shock. For one thing, we are dealing here with a situation which calls upon a host of defense mechanisms of the body, and the paralysis of any part of the neural mechanisms cannot fail to have a profound effect upon the outcome and lead to false conclusions as to the relative importance of the many factors involved. In a situation where there are such drastic changes in the circulation and metabolism, it is futile to suppose that a uniform degree of anesthesia can be maintained throughout, or even if that were possible, that its effect is only a quantitative modification of the results.

As we all know, the standardization of experimental shock has been a trying problem in nearly all investigations. In our laboratory, we were never able to standardize the production of shock by muscle trauma to the point where we could predict the fate of the animal in every instance. Dr. Walcott did, however, succeed in devising a simple method of producing hemorrhagic shock which gives 100 per cent mortality and a remarkably uniform survival time.[3] This preparation has been extremely useful to us.

The investigations which we started five years ago have covered many phases of the disturbances associated with traumatic and hemorrhagic shock.[3, 4] A summarizing table showing the average range and magnitude of the changes which were observed in dogs in severe shock after either trauma or hemorrhage has been published elsewhere.[3] In all experiments, the blood volume had been reduced by 40 to 50 per cent. In general, the disturbances are strikingly similar to those found in clinical shock.[5] In both man and dog, the reduction in blood volume in shock is approximately the same (30 to 40 per cent).[1, 6] The dog, however, displays a greater reduction in the cardiac output, but this is probably explained by the fact that, in the dog, the resting cardiac output is normally much greater in relation to size than in man.[7, 5] Also, in the dog in shock one frequently observes a 50 per cent reduction in the oxygen consumption, whereas in clinical cases of shock the oxygen consumption seldom falls below the basal level. A striking species difference appears in the peripheral resistance. According to the studies of the Bellevue Hospital group, the total peripheral resistance in cases of hemorrhage or trauma is, on the average, not far different from the normal,[5] whereas in the dog in shock, we have frequently found total peripheral resistance increased four or five times.[7] These illustrations indicate that one must exercise some caution in applying the results of animal experiments to human cases of shock. For the most part, the differences appear to be quantitative rather than qualitative.

One aim which many of us have had in the systematic study of bodily disturbances associated with shock, has been to find reliable criteria for judging the severity of the injury, the depth of the resulting shock, and for distinguishing the various stages of shock. So far as the experimental studies on traumatic and hemorrhagic shock are concerned, it seems to me that we can consider the bodily disturbances in two categories. The first category is illustrated by the change in blood volume, which gives us an idea of the severity of the injury, but is of no value as an index of the depth or stage of shock. It was perhaps a matter of luck that we happened to begin the blood volume studies on a form of experimental traumatic shock such as muscle contusion, in which the reduction in blood volume occurs suddenly at the time of injury,[1] for this demonstrated at once that the progressive character of the ensuing shock could not be ascribed to a gradual decrease in blood volume brought about, as current theories postulated, by generalized capillary leakage. The inferences which we drew from these early observations were fully confirmed by subsequent investigations of the dye disappearance rates and blood volumes in clinical cases of shock[6] and by further experiments on dogs in which it was demonstrated that the local fluid loss was always equal to or greater than the reduction in blood volume.[3]

In the second category of disturbances are those which are progressive and associated with the appearance of the clinical signs of shock.[3] These are all closely related to and result from the changes in cardiac output. As the cardiac output and total oxygen supply to the tissues diminish, the A-V oxygen difference and the circulation time through small vessels go up.[7, 9] Associated with these alterations are many evidences of the metabolic disturbance characteristic of stagnant anoxia.[10] The arterial pH and arterial CO_2 content drop, while the blood lactate and phosphate rise. These and other blood changes are now familiar to everyone who has worked on shock, and the interrelation of the metabolic changes has been clarified by contributions from several laboratories.

The question as to which of the measurements is the most accurate index of the degree or depth of shock, is rather important in evaluating the effects of various therapeutic measures. Several of my colleagues have pursued this problem but so far have not arrived at any very definite conclusions. In most of the work, we have simultaneously measured the A-V oxygen difference, the arterial CO_2, the arterial pH, and the blood lactate and phosphate. Perhaps the arterial CO_2 has been our main guide. However, even after studies on a large series of animals we are not convinced that there is a sharp limit of reduction in the arterial CO_2 below which transfusion fails to resuscitate every animal.[11, 12] The time element may be important in the outcome. It is probable that the same degree of shock as judged by the arterial CO_2, for example, may, when brought on quickly, do less damage to special systems such as the kidney and liver than when the same point is reached more slowly.

The chemical changes in the blood were much the same in hemorrhagic and traumatic shock. A survey of all our data disclosed only slight differences. For instance, for the same decrease in pH, the hemorrhaged animals showed a larger increase in phosphate than observed in traumatized dogs.[13] Also, after trauma there was a somewhat greater increase in pulmonary ventilation than after hemorrhage, which presumably accounts in part for the lower arterial CO_2 generally observed in the trauma experiments.[7] The major effects, however, are caused by the same factor, namely, a reduction in blood volume, and therefore we should not expect to see marked differences in the chemical pattern.

There were, on the other hand, differences in the superficial appearance and course of hemorrhagic and traumatic shock which aroused our curiosity early in the investigations. Evidence of pronounced central nervous depression was more striking in traumatic shock. Also, in experiments where the reduction in blood volume was the same, it was apparent that the blood pressure was, in general, maintained at a much higher level after trauma than after simple hemorrhage. This difference in the blood pressure pattern was largely abolished if the trauma experiments were done on dogs in which the spinal cord had been severed some days before.[14] Thus, there were indications of the existence of a nervous factor in the trauma experiments, but its role in shock and its effect on survival were not clear. There was, of course, the further possibility that substances released from the traumatized areas might account for some of the differences observed.

The central question was, whether or not shock and death after trauma could be ascribed entirely to the reduction in blood volume. A recent statistical analysis of the blood volume data from a large series of experiments revealed a significant difference between traumatic and hemorrhagic shock with respect to the residual blood volume associated with 50 per cent mortality (L.H.50).[15] There was, however, one flaw in the evidence, for, in the experiments on traumatic shock, ether had been used during the period when the muscles were contused,[1] whereas no anesthetic was employed in the experiments on hemorrhagic shock.[2] Wang *et al.*[15, 16] set out to repeat the experiments on hemorrhagic shock, including this time the use of ether anesthesia during the period of bleeding to make the experiments strictly comparable with the trauma experiments. The L.H.50 was the same as when no ether was employed. This evidence, incidentally, gave us much needed assurance that the brief period of ether anesthesia which had to be employed in the trauma experiments did not seriously modify the resistance of the animals to injury.

The evidence so far indicated that the afferent nerves from the traumatized region were responsible for the decreased resistance to reduction in blood volume. In order to obtain further proof, two additional series of experiments were carried out. In the first series, normal dogs were subjected to a sublethal hemorrhage, following which the central ends of the sciatic nerves were stimulated continuously.[17] In these animals, 50 per cent mortality occurred at a considerably higher residual blood volume than in dogs subjected to hemorrhage alone. Futhermore, the course of the heart rate and blood pressure and the clinical appearance of the dogs resembled more closely the picture seen in traumatic shock. In another group of dogs, all afferent nerves to the hind limbs were severed by section of the lumbo-sacral dorsal roots one to two weeks before shock was produced by muscle contusion.[18] In this series, 50 per cent mortality occurred at a significantly lower residual blood volume than in normal traumatized dogs, and furthermore, the clinical picture resembled that seen in simple hemorrhagic shock. From such evidence, there appears to be little doubt that the afferent impulses from the traumatized region play a significant role in traumatic shock. The interesting feature of these experiments is that the basis of comparison, namely, the resistance to reduction in blood volume, is objective, whereas in the attempts that other investigators have made to attack the problem, the evidence depends largely on whether or not the degree of injury has been constant from one experiment to the next.

The mechanism by which the afferent impulses from the injured region influence the course of shock and reduce the resistance to blood loss is not entirely clear. However, the evidence from the heart rate and blood pressure records suggests that the action is, at least in part, brought about through increased activity of the sympathetic nervous system, which serves to accentuate the reduction in blood flow through the tissues. Wang, Painter, and Overman have obtained additional evidence for this concept from studies of the changes in the fluorescin circulation time in shock in dogs with unilateral cervical sympathectomy.[19]

For the studies in therapy, we have almost exclusively employed dogs in hemorrhagic shock. Traumatic shock experiments are less satisfactory for this purpose, mainly because of the difficulty of controlling the leakage into the damaged area. Furthermore, we do not have a standard procedure for producing traumatic shock as is the case with hemorrhagic shock.

The first problem was to decide the stage at which the effect of transfusion should be tested. We arbitrarily selected the point at which the mean arterial blood pressure had fallen to less than 25 mm. Hg and was decreasing rapidly. The animals were therefore transfused in what was essentially the terminal stage of shock, a few minutes before *exitus*. In all instances, volume replacement with either blood or plasma more or less completely reversed the hemodynamic and chemical changes produced by the hemorrhage. It should be noted, however, that the return of the blood pressure to normal values was not in itself a reliable index of the physiological condition of the animal, for the non-surviving dogs often had high blood pressures after transfusion. Restoration of urine flow and the degree of reversal of the chemical changes in the blood were better indications of eventual survival.[11] It should be noted, also, that none of the dogs recovered immediately, on transfusion, from the lethargy which was invariably present at the time of treatment. However, the surviving animals gradually became more alert and, except for an unnatural quietness, appeared reasonably normal four hours after transfusion.

In a series of 22 dogs which were transfused in the terminal stages with their own blood, sixty per cent survived. In a second series of 16 dogs in which comparable volume replacement was carried out with plasma, three died in less than 24 hours. It should be pointed out that in spite of hemodilution, transfusion with either whole blood or plasma did not, under the circumstances of these experiments, completely restore blood volume to the control level. The deficit is explained by the blood samples taken for the various analyses. Actually, the final blood volume was 10 to 15 per cent below the initial control volume.

Certain points are brought out by comparison of the results obtained in dogs which were permanently resuscitated by transfusion with those shown by animals which finally died after volume replacement. In animals that survive, the jugular oxygen content, the arterial pH, the arterial CO_2 content, plasma inorganic phosphate, and whole blood lactate returned to control values within four hours of the time of volume replacement. The return of the various constituents is much less noticeable in animals which die within twenty-four hours. We were unable to satisfy ourselves that the absolute values of the arterial CO_2 just prior to transfusion could be utilized as a means of distinguishing dogs in which transfusion produced recovery and those which died after volume replacement. The degree to which the arterial CO_2 may be depressed without producing irreversible shock, as measured by the response to transfusion, appears to be related to the previous history and physical condition of the animal. This has been shown in a recent series in which animals were carefully selected and fed diets calculated to increase their vitamin and protein reserves.[12]

Further analysis of the data convinced us that hemorrhaged dogs which failed to recover after transfusion were in a more serious condition immediately before volume replacement than were those in which transfusions produced permanent recovery. Thus, in the terminal stages, the former had a higher whole blood lactate, a greater arterial-jugular oxygen difference, a larger increase in the plasma amino nitrogen, and, in general, a more severe acidosis than the latter group of animals. These findings suggest that when the metabolic disturbance reaches a certain stage of severity, the restoration of blood volume with either whole blood or plasma fails to resuscitate the animal permanently. Some investigators choose to term this "irreversible shock". I do not like this term, because it implies the futility of further experimental analysis. Future investigations may well bring forth effective methods of treating shock in what is now considered to be the irreversible stage.

Finally, I should like to consider, briefly, some observations on the effects of morphine in hemorrhagic and traumatic shock.[20] This problem has considerable practical interest, because, as is known, morphine has, in the past, been widely used in shock cases. Only recently, it has been pointed out that this might not be a wise routine measure. Furthermore, in the literature on shock, I have from time to time noted that investigators have claimed that no anesthesia was employed in their experiments. A careful scrutiny of the protocols, however, reveals that several doses of morphine had been administered to the animals.

In normal dogs, the intravenous injection of 2 mg. of morphine sulfate caused a sharp drop in the respiratory rate and volume, in O_2 consumption, and in blood pressure and heart rate. It is interesting to note that although the cardiac output was depressed, the calculated total peripheral resistance increased. When the same dose of morphine was given to animals in hemorrhagic shock, there was a temporary increase in the venous oxygen, arterial CO_2, and cardiac output. These evidences of improvement in the circulation did not, as far as we could tell, alter the eventual outcome or lengthen the survival time. In dogs in severe traumatic shock, the effect of morphine seemed to be only deleterious, inasmuch as it simply exaggerated the disturbances which we regard as evidences of shock. These observations call attention to the fact that even small doses of morphine may have a considerable effect on a failing circulation.

LITERATURE CITED

1. **Gregersen, M. I., & W. S. Root**
   1947. Am. J. Physiol. **148**: 98.
2. **Walcott, W. W.**
   1945. Am. J. Physiol. **143**: 254.

3. **Gregersen, M. I.**
1946. Ann. Rev. Physiol. **8**: 335.
4. **Gregersen, M. I.**
1946. Fed. Proc. **3**: 354.
5. **Richards, D. W. Jr.**
1943-44. The Harvey Lectures **39**: 217.
6. **Noble, R. P., & M. I. Gregersen**
1946. J. Clin. Invest. I **25**: 158-171; II **25**: 172-183.
7. **Root, W. S., W. W. Walcott, & M. I. Gregersen**
1947. Am. J. Physiol. **151** (1).
8. **Nickerson, J. L.**
1945. Am. J. Physiol. **144**: 429.
9. **Wang, S. C., E. E. Painter, & R. R. Overman**
1946. J. Exp. Med. **84**: 549.
10. **Root, W. S., J. B. Allison, W. H. Cole, J. H. Holmes, W. W. Walcott, & M. I. Gregersen**
1947. Am. J. Physiol. **149**: 1.
11. **Allison, J. B., W. H. Cole, W. W. Walcott, S. Gelfan, W. S. Root, & M. I. Gregersen**
To be published.
12. **Beatty, C. H., & W. L. Nastuk**
To be published.
13. **Allison, J. B., W. H. Cole, J. H. Holmes, W. W. Walcott, W. S. Root, & M. I. Gregersen**
In press.
14. **Gregersen, M. I.**
Unpublished shock report.
15. **Wang, S. C., R. R. Overman, J. W. Fertig, W. S. Root, & M. I. Gregersen**
1947. Am. J. Physiol. **148**: 164.
16. **Wang, S. C., R. R. Overman, W. S. Root, & M. I. Gregersen**
1945. Fed. Proc. **4**: 75.
17. **Overman, R. R., & S. C. Wang**
1947. Am. J. Physiol. **148**: 289-295.
18. **Wang, S. C.**
1947. Am. J. Physiol. **148**: 547-556.
19. **Wang, S. C., E. E. Painter, & R. R. Overman**
1947. Am. J. Physiol. **148**: 69.
20. **Powers, S., C. Reed, & M. I. Gregersen**
1947. Am. J. Physiol. **148**: 269-274.

VASOMOTION IN THE HEMODYNAMICS OF THE BLOOD CAPILLARY CIRCULATION

By ROBERT CHAMBERS*
Department of Biology, Washington Square College of Arts and Science,
New York University, New York, N. Y.

The blood circulation has a meaning only when we consider its hemodynamics in the capillaries. The fact that the capillaries are freely permeable to fluid, poses the problem of how a balance between outward and inward filtration can be maintained. The two main features which must be controlled are, first, an adequate venous return, and second, the periodic variations of fluid exchange between the tissues and blood.

Starling's hypothesis of hydrostatic versus colloid osmotic pressure must be supplemented by taking into account the topographical arrangement and functioning of differently structured vessels in the capillary bed. Briefly stated, the capillary bed consists of (1) preferential channels, proximally muscular (metarterioles), and coursing from the terminal arterioles to the beginning venules; (2) sphincteric offshoots which control the outflow of blood into the capillary network; and (3) the nonmuscular, true capillaries which constitute the bulk of the bed and drain into the distal portions of the channels. The preferential or thoroughfare channels are the basic structural components of the bed. The network of capillaries is built about them and is accessory to them.

This organizational pattern makes possible variations in the distribution of blood throughout a given issue and is best expressed in tissues where demands on the blood supply varies considerably whenever the tissue is in the resting or the active state.

The preferential or thoroughfare channels start directly from the terminal arterioles. Their hydrostatic pressure persists relatively high and, hence, they serve not only to transmit a positive pressure to the collecting venules leading from the bed, but also to maintain a consistent degree of outward filtration. The network of true capillaries, on the other hand, exhibits recurrent variations of pressure with concurrent shifts in its role of fluid exchange. The peculiar structure of the sphincteric precapillary offshoots, which lead into the network, and the arrangement of the postcapillaries, which lead out of it, makes for a lowering of the hydrostatic pressure to the degree that the network characteristically serves for inward filtration.

We thus have a pattern in which outward and inward filtration occur at different sites throughout the capillary bed. Two main features are controlled in this type of circulation. One is the possibility of a shift in the balance of outward and inward filtration within the capillary bed, either by keeping the blood flow predominantly in the preferential channels or by dispersing it throughout the true capillaries. The other feature is the ability of the capillary circulation, (a) to be relatively independent of pressure and of flow conditions in the systemic circulation, and (b) to permit local tissue conditions to outweigh outside influences.

---

* **Present address: Marine Biological Laboratory, Woods Hole, Mass.**

The controlling factor is vasomotion. Vasomotion is the term applied to the spontaneously recurring periods of relaxation and constriction of the thoroughfare channel and of the sphincters of its precapillary offshoots. The dilator and constrictor phases of the vasomotion alternate with one another from minute to minute, and the duration of one or the other phase varies under different physiological conditions. The vasomotion of the muscular, proximal portion of the thoroughfare channels (the metarteriole) is wavelike and exerts a milking action on the inflowing postcapillaries which lead into the distal portion of the channel. The vasomotion of the precapillary offshoots consists of a periodic opening and closing of their lumina and conditions the flow through the true capillaries. The true capillaries thus exhibit alternate periods of varying hydrostatic pressure.

The delicately balanced activity of vasomotion depends upon vasoexcitor and vasodepressor factors. Some of these, chiefly the vasodepressor, are of local tissue origin. This enables the vasomotion to act as a mechanism for locally regulating the distribution of blood and the extent and duration of inward and outward filtration in the capillary bed.

The direct action of the vasomotion affects the fluid exchange by determining the flow of blood to be either through the thoroughfare channels or by way of the true capillaries. The fluid exchange is also affected indirectly by the influence which the distribution of blood has on the rate of venous outflow from the capillary bed. For example, when vasomotion is deficient or absent, the precapillaries remain open and the propelling force of the pressure transmitted by the arterioles becomes dissipated through the numerous capillaries of the bed. This spread of flow causes a slowing in the rate of overall movement throughout the entire bed, with the result that blood tends to accumulate in the collecting venules. This induces sufficient back-pressure to favor an overall, outward filtration. On the other hand, when vasomotion is active, the flow through the capillary bed is increasingly restricted to the preferential channels, so that the propelling force of the arterial pressure is not dissipated but enhances a rapid flow in the collecting venules. The rapid venous flow favors drainage from the true capillaries in which inward filtration is thereby accentuated.

To sum up, deficiency or absence of the vasomotion shifts the balance of fluid exchange in favor of outward filtration, while enhanced vasomotion shifts the balance to inward filtration. Thus, the hemodynamic relations of fluid movement through one or another of the vascular components of the bed are of major significance in affecting fluid exchange between the blood and tissue. It follows that some types of edema are the result of disturbances in the vasomotion mechanism and do not necessarily involve alterations in the permeability of the capillary wall.

We may recapitulate the influence of vasomotion on the relative effectiveness of the hydrostatic and of the colloid osmotic pressures in the different regions of the capillary bed. Variations in vasomotion are of two sorts: (a) alterations in the rate of the intermittent caliber-changes of the thoroughfare channel; and (b) alterations in the relative duration of the constrictor and the dilator phases of the precapillary sphincters. The most effective hydrostatic pressure occurs in the thoroughfare channel, along the length of which the onward flow is fairly constant and rapid.

The colloid osmotic pressure is most effective in the network of true capillaries. However, its effectiveness in absorbing fluid from the tissue is variable. The successive periods of flow and absence of flow in the true capillaries constitute a mechanism whereby the colloid osmotic pressure becomes periodically greater than the hydrostatic pressure.

The hydrostatic pressure in the thoroughfare channels depends on the vasomotor activity of the terminal arterioles outside the capillary bed. and this activity maintains the hydrostatic pressure at a constant level despite fluctuations in the systemic blood pressure. The arterioles accomplish this by undergoing constriction when the arterial pressure rises, and by dilating when the pressure falls. This allows the capillary circulation to maintain a degree of autonomy from the pressure in the systemic circulation. Under certain conditions, this degree of autonomy is upset, *e.g.*, by the intravenous injection of fever-producing toxins. The arterioles become dilated, thereby increasing the hydrostatic pressure, but without changing the normal rhythm of the vasomotion of the thoroughfare channels. The dilated arterioles permit the blood to enter the capillary bed under an increased head of pressure. The unaltered activity of the vasomotion favors the flow through the thoroughfare channels, where the increased pressure enhances outward filtration. On the other hand, the true capillaries maintain their normal, intermittent flow with no corresponding increase of inward filtration. The result is that the overall balance is in favor of outward filtration, with consequent hemoconcentration.

It is also possible to have hemodilution with no change in the vasomotion. This occurs under circumstances when the head of pressure entering the capillary bed is significantly reduced, as, for example, immediately after the onset of bleeding in acute hemorrhage. The larger arteries and arterioles undergo vasoconstriction to accommodate the suddenly reduced blood volume. Widespread vasoconstriction occurs to compensate for the suddenly reduced blood volume. This is also reflected in the terminal arterioles, so that the effective hydrostatic pressure in the thoroughfare channels is lowered, with the consequent reduction of outward filtration. However, vasomotion remains unaltered, so that the flow in the true capillaries is undisturbed and the normal rate of inward filtration is maintained. This shifts the balance of outward to favor inward filtration and results in hemodilution.

The above phenomena illustrate imbalances in fluid exchange in the capillary bed without any disturbance in vasomotion, the imbalances resulting only from changes in the vasomotor activity of the terminal arterioles. There are also instances in which accelerated vasomotion alters the balance of fluid exchange. For example, when the loss of blood during acute hemorrhage exceeds two per cent of the body weight, the vasomotion becomes accelerated and this mechanism still further accentuates hemodilution. This is to be explained as follows. As the frequency of the vasomotion increases, the duration of the constrictor phases become prolonged, relative to the dilator phases. Concomitant with the long, constrictor phases of the precapillary sphincters, there is a progressive longer period during which the hydrostatic pressure in the true capillaries remains extremely low. As a consequence, inward filtration and tissue dehydration are accentuated. Moreover, the wavelike milking action of the enhanced vasomotion of the thoroughfare channel facilitates drainage of the true capillaries into the venous circulation.

We have been considering instances in which either the vasomotion or the state of constriction of the feeding arterioles are affected, separately. When both are affected simultaneously, edema results. A suspension of vasomotion together with a dilated state of the feeding arteriole flushes the entire bed, with consequent excessive outward and little or no inward filtration.

A list of the factors which influence vasomotion is of interest in this discussion. Enhanced vasomotion occurs with acute hemorrhage, sympathetic stimulation, intravenous administration of adrenalin, angiotonin, or adrenal cortical extract. Diminished vasomotion occurs with rise in temperature (37.5 to 41° C.), decrease in temperature by from 10 to 20° C., direct trauma, increased vital activity of the tissue (muscular exercise, secretory activity, etc.), or elaboration of vasodepressor principles following prolonged anoxia. Many agents which have been found to increase capillary circulation do not necessarily depress vasomotion. Among these are histamine, adenylic acid, adenosine, kallikrein, and acetylcholine.

There is an intimate relation between vasomotion and the rate of lymph flow in the terminal lymphatic channels. This was studied in the mesentery of the cat, dog, mouse, and rat. During the resting state of the tissue, the vasomotion is active to the extent of producing relative ischemia when little or no flow of lymph can be discerned in the terminal lymphatics. The absence of an active circulation in the bed makes the large surface furnished by the true capillaries available for inward filtration and leaves no excess of fluid in the tissues.

On the other hand, during the active state of the circulation, the vasomotion is diminished, so that an overall hyperemic flow is favored. The prolonged periods during which the precapillary sphincters remain dilated favor outward filtration in the true capillaries. This is accompanied by an appreciably increased flow of lymph in the terminal lymphatics.

REACTIONS OF PERIPHERAL BLOOD VESSELS IN EXPERIMENTAL HEMORRHAGE

By B. W. Zweifach,* Robert Chambers, R. E. Lee, and C. Hyman
Department of Biology, Washington Square College of Arts and Science,
New York University, New York, N. Y.

The problem confronting the investigator in experimental shock is a complex one, involving as it does a multiplicity of changes which occur simultaneously in the circulatory system as a whole. Since the circulatory collapse during shock is essentially peripheral in origin, a systematic inquiry into the hemodynamics of the peripheral vascular apparatus seemed to offer a direct approach to some of the more fundamental aspects of the syndrome. By focusing attention on the peripheral vascular system as a discrete organic unit with its own special physiology, it was possible to demonstrate the relationship of specific defects in this system to the broad, overall features of the syndrome.

The relative inaccessibility of the extreme peripheral portion of the vascular tree has made it the object of considerable speculation and has resulted in its being used as a convenient source of otherwise not-to-be-explained variations in the general circulation. The work of such pioneers as Krogh, Dale, Lewis, and Clark has demonstrated that the terminal ramifications of the arterial vascular tree are not merely a series of inert tubes which serve to bridge the gap between arteries and veins but represent discrete organic units whose delicately balanced activity controls the distribution of blood to the tissues. On this basis, the minutiae of the peripheral vascular apparatus, the terminal arterioles, precapillaries, capillaries, and venules are collectively referred to as the capillary bed, a system of vessels which possesses a considerable degree of independence from the circulation at large and is capable of responding selectively to tissue influences of both local and general origin.

The complexity of the peripheral vascular apparatus makes it difficult to detect changes in its component structures without direct visualization of the vessels concerned. On the whole, the chief objection to previous observational studies on the peripheral circulation has been the failure to ascertain the extent to which traumatic disturbances, incidental to preparing the tissue for observation, interfere with the normal reactivity of the vessels. It appeared essential, therefore, first, to make an intensive study of the mechanisms which integrate the distribution of blood through the capillary bed under normal conditions, and then to ascertain whether alterations in these mechanisms occur during secondary shock. Several groups of investigators[1, 2] have assigned the vascular lesion in shock to a generalized increase in capillary permeability, a hypothesis for which no supportive evidence has been forthcoming. In view of the precisely balanced character of the peripheral vascular control, a default in these circulatory adjustment mechanisms appeared to be a probable explanation of the circulatory collapse in shock. Direct microscopic observations give a continuous record which effectively mirrors the progressive impairment of the efficiency of the cardiovascular system in supplying the tissues with blood.

---

* Present address: Department of Medicine, Cornell University Medical College, and The New York Hospital, New York, N. Y.

Observations of the visceral circulation in the dog, cat, and rat following hemorrhage[3, 4] have revealed a close parallelism between the progressive changes in the terminal blood vessels and the physiologic state of the animal. The vascular reactions were sufficiently clear-cut to make it possible to resolve the hemorrhagic syndrome into two distinct sets of reactions, an initial one in which the vascular system compensates for the reduced amount of blood in circulation, and a subsequent decompensatory phase in which the factors serving to restrict the circulation are progressively undermined and, in many instances, reversed. Emphasis is placed on this latter set of changes, which appear as the shock becomes deeper and which are associated with the so-called "irreversible" state. Under conditions of prolonged generalized ischemia, such as accompanies severe hemorrhage, the terminal components of the vascular tree develop a disharmony with the rest of the circulation and act independently of the larger blood vessels to bring about a more adequate blood flow to the tissues. This decompensatory tendency of the peripheral vascular apparatus is of paramount importance in the circulatory collapse resulting from hemorrhagic procedures, since it invariably presages the development of a state refractory to blood replacement therapy.

In the present study, two types of approach were used: (a) continuous direct observations of the capillary bed were made in the shocked animal throughout the syndrome; and (b) the blood of the shocked animal was routinely examined for substances whose vascular effects could be passively transferred into test animals. This was done by observing the effects on the reactivity of the terminal arterioles of normal rats following the injection of blood samples taken from shocked animals.

METHODS

Preparation of Omentum. Dogs were anesthetized, either by sodium pentobarbital (30 mg. per kilogram), or by morphine (12 mg. per kilogram). A portion of the omentum was carefully withdrawn through a transverse incision on the left side of the dog at about the level of the umbilicus. The exposed omentum was enclosed in a rubber sheath and the incision closed by sewing the ends of the sheath to the cut edges of the peritoneum. The free end of the sheath was then cut open and the edges fastened to a specially prepared moist chamber. A small portion of the omentum was withdrawn from the open sheath and draped over a glass horse-shoe support which rested in the center of the chamber. The excess folds of the exposed omentum were covered with cotton and the entire preparation kept moist and maintained at body temperature with a constant drip of warm gelatin-Ringer's solution.

Bleeding Procedure. The dogs were then bled by successive removals, at 10 to 15 minute intervals, of 2, 1, 0.5, and 0.25 per cent of the body weight until the blood pressure was brought to the desired level. Further small bleedings or small transfusions were employed to maintain the animal within a given range of hypotension for an extended period of time. The most satisfactory method for producing a consistently reproducible hypotensive state, irreversible to transfusion, was that recommended by Wiggers.[5] This consists of a comparatively long period of moderate (50 to 60 mm. Hg), followed by a period of drastic (35 to 45 mm. Hg) hypotension. The length of the two periods varied with different dogs and with the anesthetic agent used. Our observations on the capillary bed reveal that, regardless of the type of bleeding employed to lower the blood pressure, the essential feature was the development of an inadequate peripheral blood flow over a given period of time. The more severe the stagnation, the shorter was the duration required to produce an irreversible state of shock. On the average, a 45 to 90-minute period of profound hypotension served to precipitate an irreversible type of circulatory collapse.

Vascular Criteria. The mesentery and omentum possess alternating periods of greater and lesser blood supply, during which the blood becomes periodically restricted to centrally placed preferential channels, with the majority of the capillary vessels being devoid of an active blood flow.[6, 7] The same preferential channels were the only components of the capillary bed which showed spontaneous and easily recognizable caliber changes, and were the only capillary vessels which responded to physiologic concentrations of epinephrine. It was clear that the peripheral blood vessels beyond the arterioles were not haphazardly distributed but consisted of groups of functional units, each built around a central muscular channel from which capillary side branches are given off (FIGURE 1). The immediate junctional portion of the capillary offshoot was encircled by one or two well-defined muscle cells, permitting this strategically located component to maintain a sphincteric control of the blood flow into the capillary network. For this reason, the term "precapillary sphincter" has been used to designate this component of the capillary bed.

The normal reactions of the capillary bed are regulated so as to ensure a local blood flow commensurate with the needs of the tissue. In general, it has been found that the maintenance of normal peripheral circulatory adjustments was dependent upon an equilibrium between two oppositely acting sets of factors. One, which appeared to be of both nervous and humoral origin, had a restricting influence on the capacity of the peripheral vascular bed and was characterised by a potentiating or excitor effect on the reactivity of the terminal muscular components. The second, related to cellular metabolic changes, of either local or remote origin, served to increase the amount of blood reaching the tissues and was characterized by an inhibiting or depressor effect on peripheral vascular reactivity. For example, the development of hyperemia during increased tissue activity was marked by a progressive slowing-down of the intermittent vasomotion of the terminal arterioles and precapillaries and its eventual suspension. This was then followed by a diminished responsiveness of the terminal muscular elements to epinephrine. The ensuing dilatation of these vessels, especially of the precapillary sphincters, permitted an overall flow through the entire capillary bed.

Two specific sets of reactions were selected as satisfactory criteria for reflecting deviations from the normal state of functional reactivity: (1) response of the terminal vascular elements to epinephrine; (2) variations in the character and frequency of the spontaneous, periodic changes in caliber exhibited by the terminal arterioles and the precapillary sphincters.

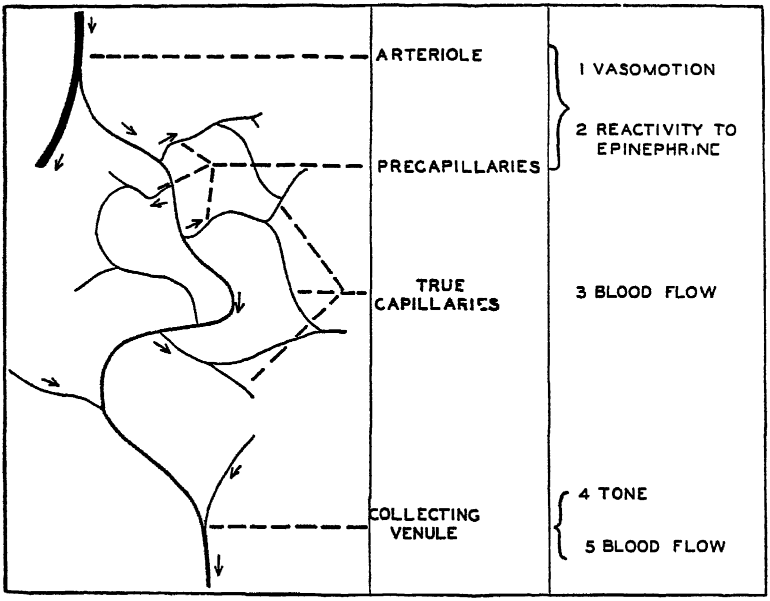


FIGURE 1. A schematic diagram of the capillary bed based on observations in the mesentery of the dog and rat. The last column lists the functional criteria upon which emphasis was placed.

The terminal muscular vessels were highly sensitive to epinephrine, responding to dilutions to which none of the larger blood vessels react. When selected arterioles were observed over an extended period of time, they showed little or no deviation from the control levels of epinephrine reactivity established at the beginning of the experiment. Persistent deflections from this basal level of epinephrine reactivity were, therefore, significant, and it was possible to distinguish two sets of changes: those associated with an enhanced responsiveness to epinephrine, and those with a decreased responsiveness.

The terminal arterioles and precapillary sphincters showed a slow intermittency of partial relaxation and constriction at intervals of about 30 seconds to 3 minutes. This intermittent activity, termed *vasomotion,* was distinct from that of the larger blood vessels and bore no apparent relation to their vasomotor movements. It serves as a mechanism for selectively restricting the peripheral blood flow and appears to be *the* mechanism which locally adjusts the peripheral blood flow to the needs of the tissue.

Following hemorrhage, the vasomotion generally became augmented and then depressed. By "augmented" is meant a condition in which the frequency of periodic contraction-relaxation cycles is increased, and in which the constrictor phase becomes increasingly predominant. By "decreased" is meant a condition in which the frequency of the contraction-relaxation cycle is diminished below normal and in which the dilator phase becomes progressively predominant.

Rat Mesoappendix Test. The presence of active humoral principles in the blood of the shocked animal was detected by noting alteration in the reactivity to epinephrine of the terminal arterioles and precapillaries, following the intravenous injection of the blood sample into normal test rats. The use of small test animals, such as rats, permitted the withdrawal at repeated intervals of blood samples sufficiently small not to overemphasize the oligemic aspects of the syndrome through the removal of excessive quantities of blood.

Previous attempts at reproducing shock effects by the injection of blood from one animal into another have relied upon blood pressure or survival as test criteria. The present approach, by relying on specific vascular changes found to occur in the shock syndrome, offered a more objective set of criteria for detection of bloodborne principles.[8]

For these tests, normal 100 to 150-gram rats were anesthetized with 3 to 4 mg. of sodium pentobarbital, and the mesoappendix was exposed for microscopic observation. An essential precaution is to keep the exposed mesentery at body temperature and constantly moistened by means of a drip of Ringer-gelatin solution. For the test, 0.5 cc. of shock serum or plasma was injected intravenously into the rat and determinations were made of the ensuing changes in the rate of blood flow, vasomotion, and response of the terminal arterioles to topically applied epinephrine. Activity of the blood samples was quantitated on the basis of the disturbance in the reaction of the vessels to epinephrine.

The minimally effective or threshold concentration of epinephrine was arrived at by testing the reaction of the arterioles and precapillaries to successively increased concentrations of epinephrine, until a concentration was found whose topical application produced a temporary, partial contraction of a selected terminal arteriole sufficient to slow down or interrupt the capillary flow for 10 to 20 seconds. Blood samples were recorded as vasoexcitor when their injection induced an increased reactivity of the terminal arterioles to epinephrine, and as vasodepressor when they brought about a decreased response. Samples were recorded as neutral when no change was detected except for a transitory speeding of blood flow due to the fluid introduced by the injection. It should be noted that the terms, vasoexcitor and vasodepressor, refer to the overall inhibiting or potentiating effects on the reactivity of the terminal blood vessels and do not refer to grosser phenomena, such as blood pressure changes or vascular dilatation.

EXPERIMENTAL RESULTS

Two Types of Peripheral Circulatory Failure. The existence of two different types of peripheral circulatory failure following hemorrhage was clearly indicated by the changes in the peripheral blood vessels. A primary differentiation was made between the effects of acute hemorrhage and those of a graded series of bleedings. The rapid withdrawal of enough blood to produce a sustained fall in blood pressure below 40 to 45 mm. Hg resulted in a mechanical upset of the peripheral blood flow of sufficient magnitude to precipitate circulatory failure before the functional integrity of the capillary bed became impaired. The peripheral circulation, under these conditions, was featured by a tremendously enhanced reactivity of the blood vessels to epinephrine and to mechanical stimuli. A different situation arose when the hypotensive state was prolonged by a series of graded small bleedings and infusions. Graded hemorrhage produced a hypotensive syndrome which was characterized by progressive deterioration of the reactivity of the functional components regulating the peripheral blood flow. When sufficiently prolonged, a stage was reached which could not be repaired by simply restoring the blood volume to normal levels.

Acute Hemorrhage. The reaction of the peripheral vascular apparatus to blood loss *per se* was essentially compensatory in nature, being an attempt to reduce the capacity of the vascular tree without drastically curtailing the blood supply to vital tissues.

It is well established that hemorrhage is accompanied by widespread vasoconstriction. Investigators who have examined the peripheral circulation in the cutaneous tissues[9, 10] agree on the development of a marked vasoconstriction resulting in complete ischemia of the skin. Page and Abell,[11] who bled dogs and cats, also observed constriction of the larger arteries and veins in the intestinal mesentery exteriorized in a specially constructed transparent chamber. We have found that the extent and intensity of vasoconstriction varied considerably in different tissues, making it difficult to generalize from observations restricted to a given tissue. When anesthetized, dogs and rats were bled acutely one to two per cent of their body weight, the first vascular change which developed was a constriction of the larger blood vessels, especially the arteries between 100 to 300 micra in diameter. By having two observers simultaneously study both the skin and visceral vessels, it was possible to demonstrate that the cutaneous vessels reacted considerably earlier to blood loss than did the visceral vessels. Moreover, the ischemia of the skin vessels was almost complete, the blood being shunted into the deeper subcutaneous plexuses. At a time when cutaneous blood flow had already been severely curtailed, the flow through the mesenteric capillary vessels remained rapid and adequate. When greater amounts of blood were withdrawn (2 to 3 per cent of body weight), the larger arteries and veins underwent a progressive narrowing to about one-half their original diameters. Despite this, the terminal arterioles and muscular venules remained open and the flow through them relatively unaffected. It was not until the depletion of blood and the accompanying vasoconstriction of the arteries were of sufficient magnitude to slow the blood flow through the peripheral vascular bed, that significant changes in the reactions of the terminal blood vessels occurred. The compensatory response of the peripheral blood vessels was initiated by the development of a hyperreactive condition in which there occurred an increasingly enhanced reactivity of the terminal muscular vessels to epinephrine, together with a five to ten-fold increase in the spontaneous vasomotor movements of the smallest arterioles and precapillaries. The development of this type of augmented reactivity served to confine the blood to the most direct, thoroughfare capillary channels, thereby reducing the amount of blood in the capillary bed and ensuring an adequate return of venous blood to the larger veins. Furthermore, the presence of a large number of open capillaries with intermittent flow made a large capillary surface available for inward filtration of fluid. This, combined with the active drainage of the capillaries which was occasioned by the periodic flushing of the bed as a result of the increased vasomotion, favored hemodilution.[12] Hematocrit readings of the blood taken during the hyperreactive period showed a decreased cellular content, the hematocrit readings falling from an average value of 42 per cent to about 30 per cent. When the blood loss was excessive (4 per cent of body weight and greater), circulatory failure developed despite the persistence of an augmented peripheral vascular reactivity. The capillary bed remained ischemic and blanched up to the point of death of the animal. The smaller arteries were completely constricted, their lumina being visible as mere lines. Only the veins and collecting venules contained blood. In these experiments, the disruption of the dynamics of the peripheral circulation was brought about by the reduced blood volume and attending vasoconstrictor changes. Blood progressively stagnated in the venous channels, the capillary bed remaining highly ischemic. Death ensued before stagnation was prolonged sufficiently to impair the functioning potentialities of the peripheral vessels. Such animals could be recovered by the expedient of the reinjection of blood withdrawn during the experiment.

Graded Hemorrhage. With this type of hemorrhage, the earlier and later stages of the syndrome exhibited striking circulatory differences. During the earlier period of the syndrome, the reactions were essentially compensatory, *viz.,* generalized vasoconstriction of the larger blood vessels, accelerated vasomotion of the terminal arterioles and precapillary sphincters, and hyperreactivity of these vessels to epinephrine. On the other hand, when the hypotension was profound and unrelieved by adequate transfusion, an antagonistic set of vascular changes progressively appeared as the shock became deeper. The initial hyperreactive, compensatory state was replaced by a hyporeactive, decompensatory state of the peripheral vascular bed. The onset of vascular decompensation was made evident by a progressive falling-off of the reactivity of the terminal arterioles to epinephrine. This tendency was associated with a slowing of the spontaneous vasomotor caliber changes in the terminal arterioles and precapillaries and eventually by a complete suspension of such movements, the vessels remaining in a state of partial relaxation. The loss of these restraining compensatory influences is indicated in TABLE 1, which summarizes the vascular changes observed under such conditions. The opening of the precapillary sphincters permitted increasing amounts of blood to be diverted into the capillary side channels, from which it failed to return to the active circulation. Circulatory collapse ensued despite the persistence of the generalized vasoconstriction of the larger blood vessels, which remained relatively unaffected throughout. The inadequate return of venous blood from the peripheral vascular apparatus to the heart left that organ incapable of maintaining an effective cardiac output and blood pressure, and a point finally was reached where a continuous circulation was impossible and respiratory and cardiac failure rapidly developed.

Criteria of Irreversibility. Blood pressure by itself was found to be an unreliable prognostic criterion. Following equivalent blood-loss, the fall in blood pressure varied considerably from animal to animal, so that no accurate correlation could be made between the blood pressure levels and the degree of impairment of the peripheral circulation. The feature which most accurately reflected the condition of the animal following hemorrhage was the ability to maintain an effective blood flow through the capillary bed. Blood loss up to 2.5 per cent of body weight had no significant effect on capillary blood flow in the omentum of the dog. With more severe blood loss, slowing of the capillary flow occurred during each bleeding and was succeeded by partial restoration of the former rate of flow. The failure of the peripheral blood flow with increased blood-loss was most evident on the venous side of the vascular tree. The relative rates of flow on the arterial and venous sides of the capillary circulation thereby provided an excellent visual prognostic guide. It was found that the blood pressure level below which the dogs had to be maintained in order to induce irreversibility, corresponded in the omentum with the tension at which the return of blood by way of the

TABLE 1

SUMMARY OF PERIPHERAL CIRCULATORY CHANGES FOLLOWING GRADED HEMORRHAGE IN THE DOG

Broken vertical lines indicate approximate onset of irreversibility to blood transfusion.

Average time in hours after initial bleeding	0	1	2	3	4	5
Blood pressure level		Moderate hypotension (70-50 mm. Hg)			Drastic hypotension (45-35 mm. Hg)	
			Dog Omentum			
Epinephrine reactivity*		hyper-reactive		transitional	hypo-reactive	
Terminal arterioles		intensely constricted		moderately constricted	relaxed	
Vasomotion		enhanced		slowed	absent	
Capillary circulation		restricted		plethoric and slow	backflow	stagnation
Flow in venules		adequate		sluggish	absent	
Tone of venules			tonic		atonic	
			*Rat Test***			
Epinephrine response		vasoexcitor			neutral	vasodepressor

* Response of terminal arterioles and precapillaries to topical application of epinephrine.

** Response of circulation in rat mesoappendix to intravenous injection of dog serum.

venules and small veins was negligible. This critical blood pressure level varied not only in different individuals, but with the anesthetic agent which was used for the experiment.[13]

During the latter stages of the hemorrhagic shock syndrome, the animal showed a repeated tendency to fail, although no blood was withdrawn. Under these conditions, the abrupt fall in blood pressure was always preceded by a complete cessation of peripheral blood flow and often by a retrograde flow of blood from the veins into the capillary bed. It would therefore appear that the continued fall in blood pressure during the hyporeactive stage was a direct result of preexisting circulatory impairment. This made the blood pressure an extremely unreliable criterion as a prognostic guide during the latter stages of the shock syndrome.

A further index of the progressive onset of an "irreversible" condition was the inability of the animal to respond to small test infusions. During the early stages of the syndrome, the infusion of small amounts of blood (2 to 3 cc. per kilogram of body weight) produced a significant rise in blood pressure (10 to 15 mm. Hg) which was frequently sustained for as long as 30 minutes. With the deepening of the shocked state, the response to test infusions became less marked and more evanescent. The refractory state of the blood pressure to infusions was paralleled by a correspondingly poor improvement in peripheral blood flow. Before irreversibility had set in, small test infusions not only produced a speeding of peripheral blood flow but caused a transient improvement in the responsiveness of the arterioles and precapillaries to epinephrine.

Analysis of Factors Leading to Irreversible Stage. Wiggers recommends that, wherever possible, a differentiation be made between initiating and perpetuating mechanisms during the shock syndrome. In so far as the capillary bed is concerned, there appear, following hemorrhage, two outstanding mechanisms. The initiating mechanism is that of widespread vasoconstriction, which follows immediately and is proportional to the blood-loss. Subsequent to this, the drastic curtailment of blood flow through the peripheral blood vessels results in the appearance of decompensatory changes antagonistic to the initial vasoconstricted state.

It is possible to emphasize one or the other mechanism by the experimental procedure employed. Thus, in acute hemorrhage, the initiating mechanism of excessive and rapid blood loss is emphasized to such a degree that circulatory collapse occurs before the development of perpetuating mechanisms. This involves a purely mechanical type of circulatory failure in which vasoconstriction of the larger blood vessels and lowered blood pressure by themselves are sufficient to produce the semistagnation of flow and pooling of blood in the capillary bed.

The hyperreactive state of the peripheral circulation progressively reverts to a subnormal condition when the animals are maintained, through the use of small infusions, in profound hypotension for a protracted period. This sets into play mechanisms, which under normal conditions would act to increase the blood flow through the tissue (diminished vasomotion with the dilator phase predominating, reduced responsiveness to epinephrine, thereby presumably reducing the response of the muscular components of the capillary bed to nervous and humoral stimuli, and the opening of numerous capillary channels to the circulation). These conditions serve to bring about a further disorganization of the peripheral circulation, a prominent feature of which is atony of the muscular components. The capillary bed now becomes an inert network of vessels with no active role in the distribution of blood and its return from the tissues. The outflow of blood through the venules becomes disproportionately reduced with respect to the arteriolar inflow, and an ever-increasing proportion of the circulating blood accumulates in the capillary bed.

There appears to be no controversy concerning the overall hemodynamics of the irreversible stage of hemorrhagic shock. Direct observation of the peripheral vascular system indicates a further intensification of the oligemic state, by sequestration of blood from the active circulation into the capillary bed, especially the collecting venules. Such a trapping of blood is also indicated by the experiments of Fine, Frank, and Seligman,[14] who produced an irreversible condition using the Lamson bottle technique and noted a persistent "taking-up" of blood from the bleed-out reservoir back into the animal with no comparable improvement in the blood pressure. Further evidence of a trapping of blood in the peripheral vascular apparatus of "irreversible" animals is seen in the appearance of the different organs following the failure to respond to the reinfusion of the blood previously withdrawn. The liver, gut, kidneys, and adrenals are highly congested, in contrast to the pale appearance of these organs following circulatory collapse induced by uncomplicated hemorrhage. It should be noted that the development of decompensatory phenomena in the latter stage of the hemorrhagic syndrome need not necessarily result in an abnormal congestion of blood in the tissues. The return of the vascular bed from a restricted to a normal state in the presence of a severely depleted blood volume is sufficient to precipitate circulatory collapse.

There remains the problem of weighing the relative importance of the factors responsible for the decompensatory vascular phenomena. It is necessary, in individual cases, to determine the extent to which the irreversible state can be attributed either to a wearing-off of vasoexcitor (VEM) influences or to the appearance of positive deleterious factors.

Fluid Replacement Therapy Following Hemorrhage. During the initial, hyperreactive stage of the hemorrhagic shock syndrome, it was possible to restore the dynamics of the capillary circulation to normal levels by the infusion of any one of the following fluids: physiological saline, 5 per cent bovine albumin, plasma, or whole blood. During the transitional stage, in which the hyperreactive aspect of the syndrome was shifting towards hyporeactivity, saline infusions became ineffective. During the latter part of this transitional stage and during the early hypo reactive stage, bovine albumin became ineffective. Finally, when hyporeactivity had definitely set in, as determined by the observed reactions of the omental vessels, recovery was no longer possible even with whole blood or plasma, irrespective of whether large amounts (up to 8 or 9 per cent of body weight) were used, or the infusion was prolonged for several hours.

The reversibility of the hemorrhagic shock syndrome by fluid infusion during the hyperreactive and early transitional stages indicates that the effects of vasoexcitor factors, which are prominent during these stages, can be counteracted by fluid therapy alone.

On the other hand, during the hyporeactive stage, the vasodepressor effects became increasingly difficult to overcome by fluid replacement. Fluid infusion during the hyporeactive stage often raised the blood pressure to approximately normal levels and set up a flow in the capillary bed, but these reactions were transitory and disappeared after the infusion was stopped (see FIGURE 2). In many instances, it was possible, by means of

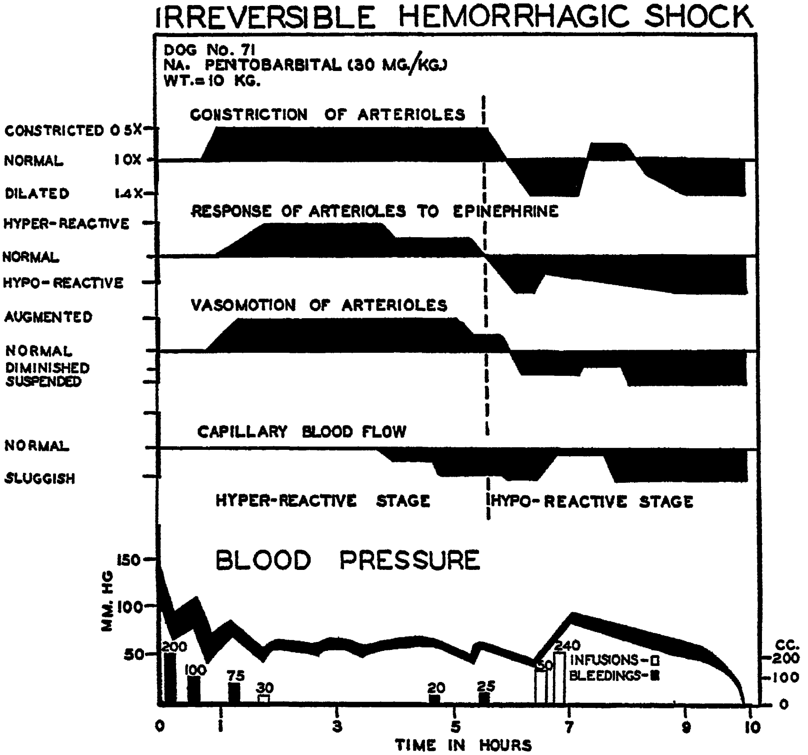


FIGURE 2. Protocol of vascular reactions in omentum of dog subjected to hemorrhagic shock and infused with whole blood during irreversible stage of syndrome. Note the failure of the vascular reactions to show an improvement comparable to the magnitude of blood pressure rise immediately following infusion of blood.

repeated infusions, to maintain an active capillary flow and to keep the blood pressure as high as 80 to 90 mm. Hg for as long as 2 to 3 hours. Despite this, there was no return of spontaneous vasomotion, the muscular vessels of the bed remained refractory to epinephrine, and atony of the arterioles and venules persisted. Under such conditions, the prolonged infusion of blood or plasma caused an abnormal dilatation of the terminal arterioles and metarterioles, indicating the loss of their normal compensatory adjustment reactions. Within 20 minutes after stopping the infusion, the capillary flow again slowed down and the dogs died shortly thereafter.

Bloodborne Principles. The relation of the vascular changes to bloodborne principles was clearly established by injecting blood samples, taken at intervals following the initial bleedings, into normal rats and observing the reaction of the terminal blood vessels in the mesentery. Blood serum or plasma from normal dogs gave no demonstrable effect other than a brief speeding of the flow through the capillaries. In the initial hyperreactive stage, the blood contained a vasoexcitor material (VEM) which was detected by its epinephrine potentiating effect, the terminal arterioles showing a five to twenty-fold increase in reactivity. With the prolongation of the profound hypotension, VEM gradually disappeared from the blood and was replaced by a vasodepressor material (VDM) whose action on the capillary bed was decompensatory, making the terminal vessels increasingly refractory to epinephrine. An interesting feature was the finding that blood taken during the transitional stage gave a neutral effect in the rat test. This suggesed the possibility that such a state might be produced by the progressive accumulation of depressor substances in blood which already contained the vasoexcitor substances acquired during the previous compensatory phase. At certain levels of concentration of each substance, the net effect in the rat mesoappendix preparation would be neutral. Attempts were therefore made to fractionate the serum obtained from dogs in the neutral phase with the anticipation that both vasoexcitor and vasodepressor substances might be demonstrated in different fractions.*

Neutral sera were fractionated by making acid-heat cuts at various pH levels. Serum samples were acidified by titrating with N. HCl to two or three pH levels, *i.e.*, pH 6.0, 4.5, and 3.5. The acidified samples were heated for 10 minutes in a boiling water bath. On cooling to room temperature, the heat-coagulated proteins were removed by centrifugation. The supernatants containing all proteins still in solution were adjusted to pH 7.2 and tested by injection into the tail vein of rats and by observing the effect on the mesoappendix circulation. By this means, it was found that samples of serum which gave a neutral effect in the rat mesoappendix tests, could be fractionated into moieties possessing pressor or depressor activity, depending upon the pH at which the acid-heat cut was made. The vasoexcitor and vasodepressor substances obtained from blood were found to be stable to heat, to be associated with definite protein fractions and non-dialyzable.

---

* Dr. M. J. Kopac was responsible for the chemical fractionation of the neutral sera samples.

These experiments indicate that in graded hemorrhage there was an early liberation into the blood of vasoexcitor substances, while during the later period of the shock syndrome there appeared a preponderance of vasodepressor material. It was during this latter period, when the vasodepressors were in the ascendancy, that the animal became irreversible to the infusion of blood or plasma.

Nature of Vasodepressor Principle. Attempts were made to reproduce the vasodepressor characteristics of shock blood by a variety of experimental procedures and by the use of known tissue extractives or drugs. Chief reliance was placed on the inhibition or depression of the response of the terminal arterioles and precapillaries to epinephrine. A summary of the effects produced by the different agents is given in TABLE 2.

TABLE 2

EFFECT OF AGENTS ON CAPILLARY BED OF RAT

Substance	Dose (per 100 gm. rat)	Effect on terminal arterioles	
		Vasomotion	Epinephrine response
Atropine sulfate	0.1 to 1.0 mg.	No effect	No effect
Acetylcholine	0.1 to 0.2 mg.	No effect	No effect
Physostigmine	0.1 to 0.2 mg.	No effect	No effect
Adenylic acid	0.1 to 0.5 mg.	Enhanced	No effect
Diphosphopyridine	1 to 3 mg.	Slow increase	No effect
Kallikrein *	5 to 20 units	Slowed	No effect
Histamine	5 to 1000 γ	No effect	Hyperreactive
Histamine	1 to 5 mg.	Depressed	Toxic
Ergotamine tartrate	10 to 100 γ	Depressed	Depressed
KCl	3 to 20 mg.	No effect	No effect
Leukotaxin **	0.5 cc.	No effect	No effect
Spreading factor (testicular) †	5 to 10 mg.	Slow depression	No effect
Shiga exotoxin ††	0.6 to 3.0 mg.	Slow depression	No effect
Cl. Welchii toxin ‡	0.05 mg.	Slow depression	No effect
Cl. Septicum toxin ‡‡	0.05 to 0.1 cc.	No effect	No effect

* Padutin (Winthrop)
** Prepared by Dr. V. Menkin; saturated solution used.
† Prepared by Dr. K. Meyer.
†† Prepared by Dr. René Du Bos; LD-50 for 25 gm. mice = 0.08 mg.
‡ Courtesy of Dr. J. Aub.
‡‡ Prepared by Dr. Colin McLeod; 75 MLD per cc.

The vascular effects of shock serum could not be reproduced in the test rat with normal serum containing free hemoglobin either in minimal traces or in amounts exceeding that occasionally encountered in shock blood. The addition of K+ ions (3 to 20 mg. KCl per 100 gram rat) to normal plasma or serum before intravenous or intramuscular injection of 100-200 mg. of KCl did not produce vasodepressor effects in the terminal arterioles and precapillaries of the test rat. Atropine, in quantities sufficient to abolish the acetylcholine effect on blood pressure (0.1 to 1.0 mg. per 100 gram rat), had no direct effect on the capillary bed. Atropinization of the test animal likewise did not abolish the vasodepressor actions of shock serum, indicating that they are probably not due to a choline derivative. Physostigmine, and acetylcholine in concentrations sufficient to lower the blood pressure (0.1 to 0.2 mg.) had no direct action on the reactivity of the vessels of the capillary bed. The injection of muscle adenylic acid (0.1 to 0.5 mg. per 100 gram rat) caused a marked slowing of peripheral blood flow and a fall in blood pressure. There was only a moderate dilatation of the terminal arterioles, with increased vasomotion and no depression of the reactivity to epinephrine. Diphosphopyridine nucleotide (1 to 3 mg. per 100 gram rat) had an action similar to, but less intense than, that of adenylic acid. A kallikrein substance derived from pancreas, Padutin (Winthrop), produced, with 5 to 20 units per 100 gram rat, a moderate fall in blood pressure and a dilatation of the muscular venules. The arteriolar vasomotion persisted except that the dilator phase was maintained for relatively longer periods than normal. Epinephrine reactivity was not affected. Minimal concentrations of 5 to 1,000 gamma per 100 gram rat of histamine produced a slight fall in blood pressure, a slowing of flow in the capillary bed, a dilatation of the venules, and a hyperreactivity of the arterioles to epinephrine, with no loss of arteriolar vasomotion. Higher concentrations (1.0 to 5.0 mg. per 100 gram rat), produced a marked fall in blood pressure, a constriction of the smaller arteries, a slowed flow, and a cessation of vasomotion with the arterioles dilated. The arterioles and precapillaries became hyperreactive to epinephrine. A variety of bacterial toxins and tissue extracts (listed in TABLE 2) were also tested, but none produced the type of vascular response characteristic of shock serum or plasma. Ergotamine tartrate, in concentrations of 10 to 100 gamma per 100 gram rat, produced an epinephrine refractory condition in the arterioles similar to that produced by shock serum.

Individual Differences in Response to Blood Loss. A wide range of individual variability in the resistance to hemorrhage is an integral feature of the syndrome. During the past several years, in our studies on vascular phenomena in shock, over 200 mongrel dogs, selected at random, were subjected to hemorrhagic procedures. The dogs could be classified under three general categories, on the basis of the amount of blood-loss tolerated. About 12 per cent of the animals withstood considerable blood-loss, as high as 6.0 to 6.5 per cent of body weight, before showing symptoms of shock. At the other extreme was a somewhat larger group, about 22 per cent of the animals, who were precipitated into shock by comparatively small bleedings, 2 to 3 per cent of body weight. The majority of the animals, 67 per cent, were intermediate between these, requiring about 4 to 5 per cent blood loss to precipitate shock.

The behavior of the animals in this respect was reflected by their tendency towards compensatory or decompensatory vascular responses. Thus, the resistant group of animals went into fatal circulatory collapse with compensatory reactions, such as vasoconstriction, increased frequency of intermittent caliber changes of terminal arterioles, and hyperreactivity to epinephrine persisting until shortly before death. On the other hand, dogs susceptible to hemorrhage showed only moderate or no compensatory vascular changes. Hyporeactivity developed early and was followed by a progressive deterioration of peripheral blood flow. Circulatory collapse occurred following the progressive suspension of spontaneous vasomotor caliber changes and ultimate unresponsiveness of the arterioles to epinephrine. The large majority of animals showed a blend of initial compensatory and subsequent decompensatory vascular phenomena.

The factors leading to a collapse of the circulation can best be considered by noting that, basically, shock develops when there no longer is a delicate equilibrium between the amount of blood in actual circulation and the capacity of the vascular tree. Two fundamentally different types of disturbances can initiate and contribute to an unbalanced hemodynamic state. On the one hand, there are those factors directly affecting the blood volume, such as hemorrhage, or the loss of plasma into traumatized tissues, or dehydration. On the other hand are indirect factors which unduly increase the vascular space through which the blood is being circulated and result in an inadequate return of blood from the peripheral vascular apparatus to the heart, thereby reducing the effective circulating blood volume. Thus, with regard to the basic mechanisms underlying the collapse, it is possible to have two widely divergent types of shock. At one extreme would be uncomplicated hemorrhage where the reduction in blood volume is solely responsible for the ensuing circulatory failure. At the other extreme would be the vasodilatation and visceral congestion frequently observed following extensive trauma[15] with little or no change in the actual blood volume.

Circulatory collapse attributable solely to either of these extremes is rare. The majority of shocked states involve a multiplicity of factors. This is especially true of laboratory experiments on animals where anesthesia is a contributing factor. Perhaps the most important outcome of the shock work done during the past several years has been the realization that no one set of circumstances can serve as an explanation for the changes observed under different conditions. There exists, rather, a variety of shocked states in which a variety of factors participate.

Effect of Anesthetic Agents. Experiments designed to investigate shock in the laboratory have regularly utilized anesthesia as a means of immobilizing the animal for the technical procedure. A tabulation of the circulatory data showed striking differences not only between animals bled under different anesthetic agents but also between anesthetized and unanesthetized dogs (TABLE 3). Unanesthetized dogs required considerably larger bleedings to keep them within hypotensive levels than either

TABLE 3

PREDISPOSING ACTION OF ANESTHESIA ON RESPONSE TO HEMORRHAGE IN THE DOG

Anesthetic agent	Blood loss tolerated % body wt.	Critical * B.P. level mm. Hg	Irreversibility produced by mm. Hg × hrs.	Reactivity of arterioles to epinephrine	
				before bleeding	at death
None	5.5 - 6.5	30	—	1:12 M**	1:10 M
Morphine 2 mg./Kg.	4.5 - 5.0	40	35 × 1.5	1:8 M	1:2 M
Nembutal 30 mg./Kg.	3.5 - 4.5	45	40 × 0.75	1:4 M	1:1 M

* = that below which capillary circulation is drastically curtailed.
** = minimal effective concentration of epinephrine producing threshold reaction of arterioles, expressed as 1 part in x million (1:x).

morphinized or nembutalized dogs (FIGURE 3). Unless kept below 50 mm. Hg by repeated small bleedings, dogs with no anesthesia showed a spontaneous tendency to climb out of shock. This was in marked contrast to nembutalized dogs which were in deep shock when maintained at the 45 mm. Hg level. Unanesthetized dogs could be maintained at blood pressure levels of 30-35 mm. Hg for as long as three to four hours. At the end of this period, these dogs were not only reversible but responded

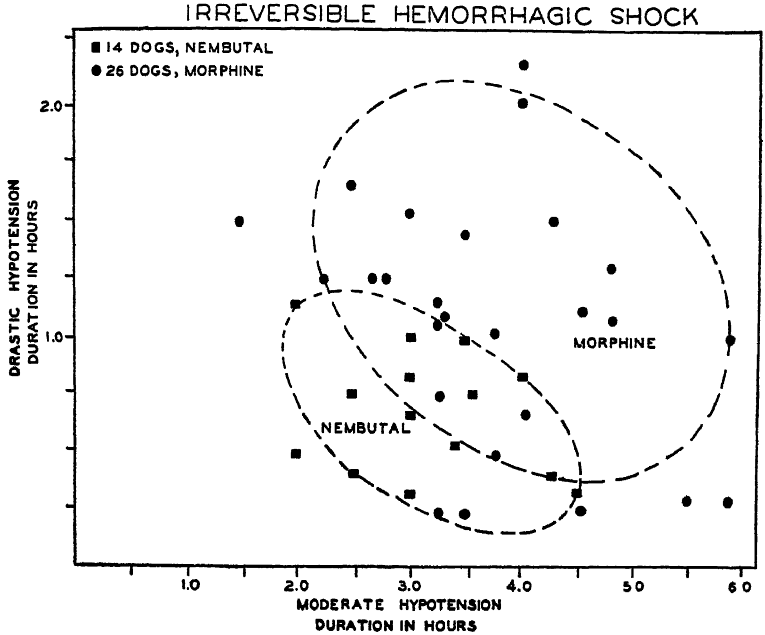


FIGURE 3. Comparison of degree of hypotension necessary to produce irreversible state in dogs under Nembutal (30 mg./Kg.) or morphine sulfate (2 mg./Kg.) anesthesia. The nembutalized dogs were more susceptible to shock and required a less drastic hypotension to precipitate irreversibility.

to the infusion of as little as 5 to 7 cc. of whole blood per kilogram of body weight. The femoral incision (for blood pressure recording and bleedings) was then closed, and the dogs were able to walk about immediately on release and usually drank water copiously.

In comparison with unanesthetized dogs, most of the anesthetic agents shifted the tendency towards decompensatory vascular reactions. The development following uncomplicated hemorrhage, of a state irreversible to transfusion appeared to be restricted to anesthetized animals. The depressant action of anesthesia served to predispose the animal to the development of a vasodepressed state of the peripheral circulation. The choice of anesthesia is therefore an important factor in studies on experimental hemorrhage.

BIBLIOGRAPHY

1. **Freeman, N. E.**
   1943. Shock. In: *Burns, Shock, Wound Healing and Vascular Injuries.* Saunders. Philadelphia.
2. **Moon, V. H.**
   1942. Shock, its Dynamics, Occurrence and Management. Lea & Febiger. Philadelphia.
3. **Zweifach, B. W., R. E. Lee, C. Hyman, & R. Chambers**
   1944. Ann. Surg. **120:** 232.
4. **Zweifach, B. W., B. E. Lowenstein, & R. Chambers**
   1944. Am. J. Physiol. **142:** 80.
5. **Wiggers, C. J., & J. M. Werle**
   1942. Proc. Soc. Exp. Biol. & Med. **49:** 604.
6. **Chambers, R., & B. W. Zweifach**
   1944. Am. J. Anat. **75:** 173.
7. **Chambers, R., & B. W. Zweifach**
   1946. Ann. N. Y. Acad. Sci. **46**(8): 683.
8. **Chambers, R., B. W. Zweifach, B. E. Lowenstein, & R. E. Lee**
   1944. Proc. Soc. Exp. Biol. & Med. **56:** 127.
9. **Mann, F. C.**
   1915. Surg. Gynec. Obst. **21:** 430.
10. **Meek, W. J., & J. A. E. Eyster**
    1921. Am. J. Physiol. **56:** 1.
11. **Page, I. H., & R. G. Abell**
    1943. J. Exp. Med. **77:** 215.
12. **Chambers, R.**
    1948. Ann. N. Y. Acad. Sci. **49** (4): 549.
13. **Zweifach, B. W., S. G. Hershey, E. A. Rovenstine, R. E. Lee, & R. Chambers**
    1945. Surgery **18:** 48.
14. **Fine, J., A. M. Seligman, & H. A. Frank**
    1946. Personal communication.
15. **Zweifach, B. W., R. G. Abell, R. Chambers, & G. H. A. Clowes**
    1945. Surg. Gynec. Obst. **80:** 593.

HEPATO-RENAL FACTORS IN CIRCULATORY HOMEOSTASIS:

III. THE INFLUENCE OF HUMORAL FACTORS OF HEPATO-RENAL ORIGIN ON THE VASCULAR REACTIONS TO HEMORRHAGE*

By EPHRAIM SHORR, B. W. ZWEIFACH, AND ROBERT F. FURCHGOTT

Department of Medicine, Cornell University Medical College, and
The New York Hospital, New York, N. Y.

A clarification of the vascular changes which occur in experimentally produced shock provided the stimulus for the present inquiry[1] into metabolic disturbances which could be related to the vascular dysfunctions observed. Direct visualization of the peripheral blood vessels in the omentum and mesentery has demonstrated that the shock syndrome consists of two consecutive stages, an initial compensatory phase, apparently related to a reduction in blood volume, and a subsequent decompensatory stage, related to the period of inadequate peripheral blood flow.[2]

The reaction of the peripheral vascular apparatus to blood loss *per se* was found to be essentially *compensatory* in nature; that is, an attempt to reduce the capacity of the vascular tree without drastically curtailing the blood supply to vital tissues. The compensatory response is characterized by a hyperreactive condition of the peripheral blood vessels, as evidenced by an increasingly enhanced reactivity of the terminal muscular vessels to epinephrine and a five to ten-fold increase in the spontaneous vasomotor movements of the smallest arterioles and the precapillaries.[3] This type of heightened reactivity persists even when the blood loss is sufficient to precipitate the collapse of the animal. The resulting capillary ischemia and the restriction of peripheral blood flow to the most direct, thoroughfare channels serve to maintain an active venous return of blood from the tissues until shortly before death.

On the other hand, with the development of an irreversible type of shock following prolonged drastic hypotension, a set of new vascular changes progressively appear and finally disrupt the peripheral circulation. These are *decompensatory* in nature and are initially made evident by a progressive falling-off of epinephrine reactivity and a slowing and, finally, complete cessation of spontaneous vasomotor caliber changes in the terminal arterioles. The loss of these restraining compensatory influences on the peripheral circulation, results in the diversion of increasing amounts of blood into the capillary side channels from which it fails to return to the active circulation. Decompensatory changes of a vasodepressor nature were consistently obtained in animals which were subjected to prolonged hemorrhagic hypotension or to traumatic procedures.[4]

---

* The work described in this paper was begun under a contract recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Cornell University Medical College, and continued under grants from The Josiah Macy, Jr. Foundation and the Eli Lilly Company. We wish to express our thanks to Doctor Silvio Baez of The Asuncion Medical College, Ascuncion, Paraguay, for the many ways in which he has assisted in these studies.

The relation of the vascular changes to blood-borne principles was clearly established by injecting blood samples taken at intervals during the shock syndrome into normal rats and observing the reaction of the terminal blood vessels in the mesentery.[5] In the initial, hyperreactive stage of shock, the blood contained a vasoexcitor material (*VEM*) which was detected by its epinephrine potentiating effect, the terminal arterioles showing up to twenty-fold increase in reactivity. With the prolongation of profound hypotension, VEM gradually disappeared from the blood and was replaced by a vasodepressor material (*VDM*) whose action on the terminal vessels was decompensatory, making them increasingly refractory to epinephrine.

The significance of these observations lies in their demonstration of the constant participation, in the shock syndrome, of humoral vasotropic principles with specific actions on the terminal vascular bed of such a character as to suggest that they are causally related, both to the initial compensatory vascular reactions to blood loss, in the case of VEM, and, in the case of VDM, to the progressive vascular deterioration which eventually results from the prolongation of drastic hypotension.

The mechanisms responsible for the occurrence of these humoral principles remained obscure until experimental conditions were realized which revealed their sites of origin in specific tissues in the shocked animal, as well as the environmental and cellular factors responsible for their formation and destruction.[1] These studies involved both *in vivo* and *in vitro* procedures, whose correlation was made possible by the rat mesoappendix test of Zweifach and Chambers[5] for assaying VEM and VDM. The present report presents an analysis of the conditions and of the tissues to which the genesis of these vasotropic factors in hemorrhagic shock could be related.

SITES OF ORIGIN OF VASOEXCITOR AND VASODEPRESSOR MATERIAL

The first group of experiments was designed to trace to their tissue origins the VEM and VDM which appear in the blood stream during the hyper- and hyporeactive stages of hemorrhagic shock. Their appearance in a definite sequence during the development of the syndrome suggested that these factors might be products of the deranged metabolism of specific organ systems rather than the manifestation of a general disturbance in tissue metabolism.

Methods. Shock was induced in dogs by graded hemorrhage according to the method suggested by Wiggers[6] for producing a state irreversible to blood replacement therapy. The initial bleedings were such as to reduce blood pressure to levels of 70-90 mm. Hg and to maintain these levels for 90-120 minutes. Thereafter, the animals were maintained, by further bleeding, in extreme hypotension at blood pressures below 60 mm. for periods of from 90 to 250 minutes. Except where otherwise noted, the animals were under sodium pentobarbital anesthesia, the usual dose being 30 mg. per kilogram. Blood samples taken at intervals during the course of the shock syndrome provided the basis for differentiating the hyperreactive or compensatory phase from the hyporeactive, decompensatory or irreversible stage. The temporal association of humoral VEM with vascular hyperreactivity and of VDM with hyporeactivity had been previously established by Zweifach, Chambers, and their associates.[2] At appropriate times in the shock syndrome, the animals were sacrificed and a variety of tissues removed for study. Thin slices were made of heart, spleen, liver, and kidney as for micro-respiration studies. Sheets of smooth muscle were prepared from the small intestine and individual skeletal muscle fibers dissected from the thigh muscles *in situ*. The tissues were immediately placed in five times their weight of chilled physiological saline, which was constantly agitated for 5 to 7 minutes by a stream of oxygen to facilitate the extraction of diffusible material. Supernatants were centrifuged clear of debris and 0.5 cc. injected for bioassay into the tail vein of 100-125 gm. rats. The presence of VEM was manifested by an increase, and that of VDM by a decrease, in the reactivity of the terminal vascular bed of the meso-appendix to the topical application of epinephrine. Both factors were graded in terms of the intensity and duration of these alterations in vascular behavior.

Experimental Results:

Vasotropic Content of Tissues During the Hyporeactive Phase. In this group of experiments, tissues were removed for bioassay during the hyperreactive stage of hemorrhagic shock, usually 90 to 120 minutes after the initial bleeding, by which time significant amounts of VEM had appeared in the blood stream. The only tissue to which its genesis could be related was the kidney, saline extracts of which invariably contained considerable amounts of VEM, whose vascular effects in the test rat were identical with those induced by humoral VEM. The concentration of VEM in the renal parenchyma appeared to be at least as great as, and probably greater than, that in blood, to judge from the relative activities of undiluted serum and of the renal extracts which represented a five-fold dilution of the tissue content. Saline extracts of liver, spleen, cardiac and smooth muscle were either neutral, or, in the case of liver and spleen, occasionally contained small amounts of VEM, attributable to the blood retained in these organs (FIGURE 1). Skeletal muscle washes were also neutral or contained a small amount of bloodborne VEM.

Vasotropic Content of Tissues During the Hyporeactive Phase. In this group of dogs, after an initial hyperreactive phase lasting 90-120 minutes, the blood pressure was further reduced by additional bleedings, and maintained at around 40 mm. for periods of 90-250 minutes. VEM gradually disappeared from the blood stream and was replaced by increasing amounts of VDM. When the hyporeactive stage was well established, as indicated by the concentration of humoral VDM, tissues

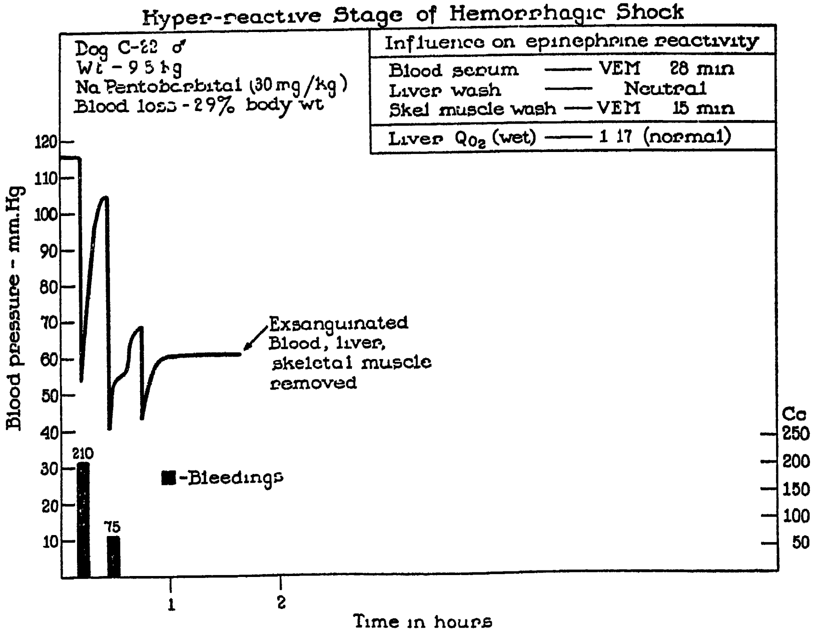
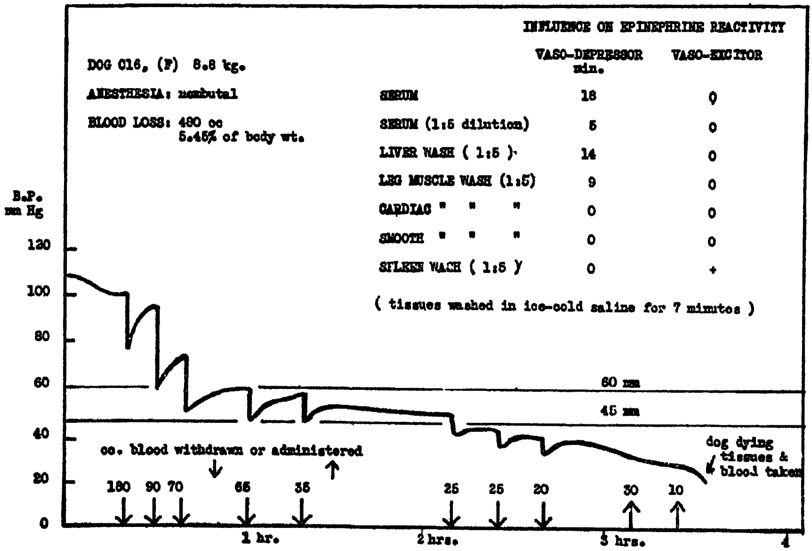


FIGURE 1. Protocol chart of dog sacrificed during hyperreactive, reversible stage of hemorrhagic shock. The rat test data in the rectangle show the failure of VDM to appear in either the tissues or the blood during this stage. Blood VEM activity was high. The oxygen consumption of the liver fell within the normal range.

were removed and extracted for bioassay (FIGURE 2). Saline extracts of spleen, cardiac and smooth muscle were uniformly neutral. Those from liver invariably contained considerable amounts of VDM whose vascular effects on the rat mesoappendix were indistinguishable from those induced by humoral VDM. The concentration of VDM in the liver was always greater than that in serum obtained simultaneously and similarly diluted. Skeletal muscle washes contained small to moderate amounts of VDM, the concentrations varying directly with the duration of the hypotension.

As in the hyperreactive stage, kidney extracts contained VEM, but in smaller amounts. In experiments in which the period of drastic hypotension was prolonged by transfusions, little or no VEM was present in the kidney washes. When slices from such kidneys were then incubated anaerobically, they proved to be incapable of elaborating VEM, in contrast with normal kidneys. The significance of this progressive reduction in VEM content of the kidney with the prolongation of drastic hypotension and of the eventual loss of the capacity to form VEM on anaerobic incubation, will be discussed later.

These experiments provided presumptive evidence for the renal origin of the humoral VEM present in the hyperreactive stage of hemorrhagic shock, and for the genesis in liver, and to a lesser degree in skeletal muscle, of the humoral VDM characteristic of the hyporeactive or decompensatory phase. However, it seemed desirable to devise experimental procedures which might provide additional and, if possible, more specific support for these inferences.



INFLUENCE ON EPINEPHRINE REACTIVITY

	VASO-DEPRESSOR min.	VASO-EXCITOR
SERUM	18	0
SERUM (1:5 dilution)	5	0
LIVER WASH (1:5)	14	0
LEG MUSCLE WASH (1:5)	9	0
CARDIAC " " "	0	0
SMOOTH " " "	0	0
SPLEEN WASH (1:5)	0	+

(tissues washed in ice-cold saline for 7 minutes)

FIGURE 2. Protocol chart of dog in hyporeactive, irreversible stage of hemorrhagic shock. In upper right hand corner are results of rat mesoappendix test assay of blood and of saline washes of different tissues removed at *exitus*. Only liver and skeletal muscle contained VDM, a considerably higher concentration being present in the liver. In comparing blood and liver VDM levels, it should be noted that the liver wash represents a 1:5 dilution of the VDM originally present in the liver.

Influence of Duration of the Hyperreactive Phase on Vasotropic Content of Tissues. In the previous studies of the hyperreactive phase, the tissues were removed for bioassay approximately 90-120 minutes after the initial bleeding. The absence of VDM in the liver wash at these times might conceivably have been due to the brief duration of the hyperreactive phase. To resolve this uncertainty, shock was induced in unanesthetized dogs by graded hemorrhage, so as to achieve the same degree and duration of drastic hypotension which leads to hyporeactive shock in anesthetized animals. Under these circumstances, unanesthetized animals remain in the hyperreactive stage until *exitus*, and are recoverable by transfusion. Significant amounts of VEM persist in the blood stream throughout. The tissues were removed at *exitus*, approximately 4 hours after the initial bleeding. The results of their bioassay were similar to those in anesthetized dogs in which the hyperreactive stage was of shorter duration. Despite the prolongation of the period of drastic hypotension, the liver was devoid of VDM. The kidneys contained large amounts of VEM. Minimal amounts of VDM were present in skeletal muscle. Thus, as far as liver VDM is concerned, its formation appears to be restricted to the hyporeactive phase.

Time of VDM Formation in Liver During Hemorrhagic Shock, as Revealed by Serial Liver Biopsies. Even more direct evidence of the temporal relationship between the formation of VDM by liver and the hyporeactive stage of hemorrhagic shock, was derived from experiments in which serial liver biopsies were obtained for bioassay for vasotropic factors during the hyper- and hyporeactive stages of hemorrhagic shock in anesthetized animals(TABLE 1). Bioassays were also carried out on blood samples obtained simultaneously with liver biopsies. During the hyperreactive phase, when the blood contained VEM, the saline extract of the liver was neutral. The second biopsy obtained at the termination of the hyperreactive phase as indicated by the neutral reaction of blood, revealed the onset of VDM formation in the liver. At *exitus*, after the animal had been in drastic hypotension for approximately 3½ hours, both liver extracts and blood contained large amounts of VDM. These results demonstrated that VDM formation by the liver does not take place until the hyperreactive phase is terminated, and that VDM occurs in the liver prior to its appearance in the blood stream.

TABLE 1

RELATION OF THE STAGE OF HEMORRHAGIC SHOCK TO HEPATIC VDM FORMATION AND INACTIVATION CAPACITY AS REVEALED BY SERIAL LIVER BIOPSIES *

Sample No.	Phase of shock	Duration of Hypotension	VDM		Inactivation of VDM by liver in O_2†
			Blood **	Liver biopsy	
1	Control	0	0	+	15/0 = 100%
2	Hyperreactive	<75 mm. for 25 minutes	VEM 16′	0	15/3 = 80%
3	Transitional	<60 " " 100 "	0	++	15/4 = 73%
4	Hyporeactive	<60 " " 185 "	++	+++	15/9 = <40%

* Sodium pentobarbital anesthesia, 30 mg./kg.

** The blood samples were taken simultaneously with the liver biopsies.

† The percent inactivation is expressed as a fraction, with the original VDM activity of the sample as the numerator and the residual activity after 2 hours incubation as the denominator.

The temporal association of liver VDM with the hyporeactive stage of shock having been established, experiments were then carried out to ascertain the dependence of the hyperreactive phase on renal VEM formation and its release into the circulation. These entailed the introduction of one modification in the experimental procedure; the exclusion of the kidneys from the circulation during the shock syndrome.

The Hemorrhagic Shock Syndrome in Arenal Animals. When both kidneys were tied off just prior to the induction of hemorrhagic shock, the sequence of events was strikingly different from that which obtains in anesthetized dogs with intact renal circulation. The hyperreactive stage was abolished, as evidenced by the absence of humoral VEM, and the hyporeactive stage set in with unusual rapidity. The concentration of humoral VDM rose to unusually high levels at *exitus,* and very considerable amounts of VDM were present in the liver wash.

Since hemorrhagic shock in unanesthetized animals is characterized by the persistence of hyperreactivity until *exitus,* the induction of hemorrhagic shock in such animals should provide a crucial test of the concept that renal VEM is essential for the development and maintenance of the compensatory hyperreactive vascular response to hemorrhage. For this purpose, one kidney was removed and the other exteriorized in the flank. After a suitable recovery period, the exteriorized renal pedicle was ligated under procaine anesthesia just prior to bleeding. In contrast to the findings in unanesthetized animals with intact kidneys, no VEM appeared in the blood stream and, within a relatively short time of the initial bleeding, humoral VDM appeared in increasing concentrations, indicating the rapid development of the hyporeactive stage. Unusually large amounts of VDM were present in saline extracts of liver removed after 5½ hours of drastic hypotension. Irreversibility was demonstrated by the failure to respond to the complete replacement of the blood lost. No satisfactory explanation has hitherto been provided for the persistence of hyperreactivity and reversibility in unanesthetized animals, in the face of prolonged hypotension drastic enough to regularly induce hyporeactivity and irreversibility in anesthetized animals. These experiments indicate that, in the absence of anesthesia, there persists a renal blood flow of sufficient magnitude to deliver renal VEM continuously into the general circulation at hypotensive levels which abolish the renal circulation in the anesthetized animal.

Recapitulation. In hemorrhagic shock, the hyperreactive phase which is associated with the presence of VEM in the blood stream is dependent on the continued formation of VEM in the kidney and on its delivery to the circulation. The drastic curtailment of the renal circulation which results from the induction of profound hypotension, blocks the release of VEM into the circulation and leads to the abolition of the hyperreactive stage. The exclusion of the kidney from the circulation by ligation prevents the appearance of a hyperreactive stage and leads to the rapid onset of hyporeactivity even in unanesthetized animals. The development of the hyporeactive stage is dependent upon the formation of VDM in the liver. This does not take place while the vascular behavior is dominated by humoral VEM, but as VEM disappears from the blood stream, VDM formation is initiated in the liver, from which it is delivered into the blood stream in increasing amounts as the period of drastic hypotension is prolonged. One additional source of VDM appears to be skeletal muscle, but the amounts available from this source in hemorrhagic shock are much smaller than those from the liver.

There remained to be determined the mechanisms which lead to the formation of these vasotropic factors during hemorrhagic shock, the most probable being the reduced oxygen tensions resulting from the reduction in blood flow to the tissues in question. Since these mechanisms can be more directly investigated by *in vitro* procedures, such studies were carried out with a variety of tissues previously studied.

MODE OF ORIGIN OF VASOEXCITOR AND VASODEPRESSOR MATERIAL

The following tissues: liver, kidney, spleen, brain cortex, cardiac, smooth and skeletal muscle, were obtained from normal dogs and rabbits and prepared for *in vitro* studies in the manner previously described, except for the brain cortex, which was minced. The tissues were then incubated at 37.5°C in normal serum, or, more generally, in Krebs bicarbonate medium under an atmosphere of 95% N_2-5% CO_2 at pH 7.4. Control observations were carried out in an atmosphere of 95% O_2-5% CO_2. After varying periods of incubation, the medium was freed of debris by centrifugation and 0.5 cc. was injected into the tail vein of the rat for bioassay by the mesoappendix test.

Experimental Results. *In vitro* results were in complete agreement with tissue extracts from shocked animals as regards the site of origin of VEM and VDM. Tissue anoxia was found to be the condition necessary for their formation, neither factor occurring in any tissue during aerobic incubation. The supernatants from spleen, brain cortex, cardiac and smooth muscle were entirely devoid of vasotropic activity (TABLE 2).

TABLE 2

In vitro PRODUCTION OF VDM AND VEM ON ANAEROBIC INCUBATION AT 37.5° C.

	VDM (2 hours)	VEM (1 hour)
Liver	++++	0
Skeletal muscle	+	0
Kidney	0	++++
Smooth muscle	0	0
Cardiac muscle	0	0
Spleen	0	0
Brain cortex	0	0
Blood serum or plasma	0	0

In vitro *Formation of VEM.* VEM formation was restricted to the anaerobic kidney and confined to the cortical portion. Its rate of formation was rapid, a considerable concentration being reached in 15 minutes, and peak concentrations at 45 to 60 minutes. With a further prolongation of the period of anaerobic incubation, VEM was found to disappear from the medium, none being detected at the end of 2-3 hours of anaerobiosis (FIGURE 3). Apparently the prolongation of anaerobiosis initiates other processes, presumably of a proteolytic nature, which eventually effect the destruction of VEM which has been formed during the initial period of anaerobiosis. In view of the possibility that the eventual disappearance of VEM on prolonged incubation might, however, have been due to the depletion of substrate necessary for its formation, an additional period of anaerobiosis was carried out in the presence of fresh normal plasma, without any further appearance of VEM.

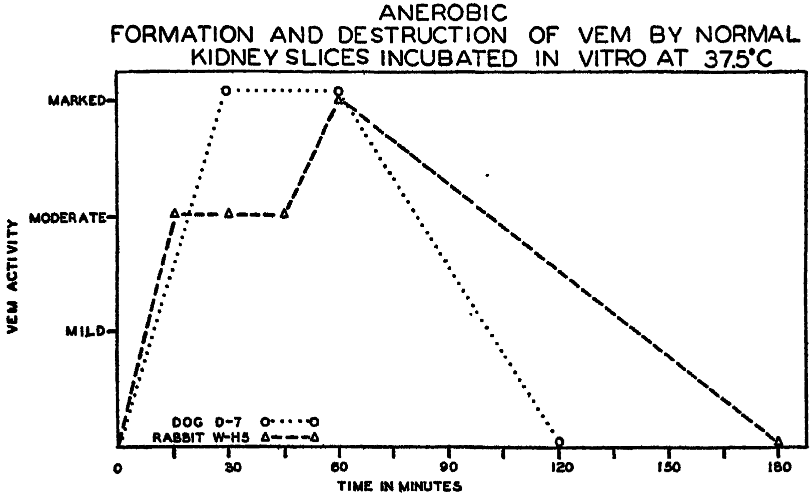


FIGURE 3. Tissue slices were prepared immediately after removal of the kidney from normal anesthetized animals. They were then washed free of preformed VEM prior to incubation in a series of flasks in Krebs-bicarbonate medium, pH 7.4, in an atmosphere of 95 per cent nitrogen —5 per cent carbon dioxide. Each point on the graph represents the VEM content of the medium from an individual flask whose anaerobic incubation was terminated at the time noted. The rapid initial formation, and eventual disappearance of VEM on prolongation of anaerobiosis, are comparable to the successive changes in renal VEM content during the renal anoxia of a prolonged hyporeactive stage of hemorrhagic shock; with this difference, that the disappearance of renal VEM takes place more slowly *in vivo*, because of the less complete renal anoxia.

The VEM which resulted from anaerobic incubation of normal kidney tissue was identical in its action on the vascular bed of the rat mesoappendix with that of humoral VEM obtained during the hyperreactive stage of hemorrhagic shock and of saline extracts of kidneys of shocked animals.[*]

In vitro *Formation of VDM.* Liver and skeletal muscle invariably produced VDM similar, in action on the vascular bed, to VDM appearing in blood during the hyporeactive phase of shock and in extracts of liver and skeletal muscle removed from shocked animals. The amounts formed during anaerobiosis were of the same order of magnitude as those present in washes of shock tissues. The rate of formation was much greater in liver than in skeletal muscle. With liver, no VDM was detectable for about 20-30 minutes. Thereafter, the amounts increased quite rapidly and progressively throughout a 3-hour exposure. Larger amounts appeared after incubation with serum than in Krebs bicarbonate. Apparently, complete anaerobiosis was not essential, since VDM was also formed by liver *in vitro* at oxygen tensions of 5-10 per cent, conditions more comparable to those prevailing in the liver during the hyporeactive stage of shock. Only trivial amounts were formed by liver brei, indicating the necessity of cellular integrity for its maximal production.

VDM formation during anaerobic incubation of skeletal muscle proceeded at a much slower rate than with liver. In view of the well established fact that the severity of the shock syndrome is increased by environmental conditions which raise the temperature of the limbs, it was of interest to investigate the influence of temperature on the rate of VDM formation by skeletal muscle. When the temperature during incubation was varied from 21° to 41° C., there was a progressive and significant increase in VDM formation with the progressive elevation of the temperature. This phenomenon is of particular significance for traumatic shock, in which the VDM produced by the muscles of the tourniqueted limbs has been shown to be an important contributing factor to the development of the hyporeactivity,[1] and may be responsible for the more profound reaction to equivalent reductions in blood volume manifested by animals in traumatic as compared with hemorrhagic shock.

The possibility that VDM and VEM were products of bacterial metabolism was investigated and excluded.

MODE OF DESTRUCTION OF VASOEXCITOR AND VASODEPRESSOR MATERIAL

The disappearance of renal VEM on prolonged anaerobic incubation suggested that the kidney possessed mechanisms not only for its formation but also for its inactivation. The possibility that the organism likewise possessed a mechanism for the destruction of VDM was suggested by the observation that VDM injected into the normal test rat disappears from the circulation usually within 10-30 minutes, depending upon its initial concentration; whereas, once significant amounts of VDM have appeared in shock, this principle persists in the circulation until death and is subject to only temporary dilution by transfusions. The *in vitro* approach proved most suitable for ascertaining whether such an inactivating mechanism existed for VDM, and whether it was a general tissue phenomenon or limited to specific organs.

Experimental Results. Tissues from normal animals were incubated aerobically at 37.5° C. with VDM obtained from various sources, including humoral, hepatic, and skeletal VDM from animals in hyporeactive shock and VDM resulting from anaerobic incubation of liver and skeletal muscle *in vitro*. A similar survey was made with VEM derived from the blood and kidney of shocked animals and from the anaerobic incubation of normal kidney tissue.

Inactivation of VDM. Spleen, kidney, cardiac, smooth and skeletal muscle proved incapable of destroying VDM (TABLE 3).

TABLE 3

In vitro INACTIVATION OF VDM AND VEM ON AEROBIC INCUBATION FOR TWO HOURS AT 37 5° C

	VDM	VEM
Liver	++++	++
Skeletal muscle	0	0
Kidney	0	++++
Cardiac muscle	0	0
Smooth muscle	0	0
Spleen	0	0
Blood serum or plasma	0	0

On the other hand, normal liver slices invariably destroyed VDM from whatever source obtained, in the course of a 2 to 3-hour aerobic incubation. Thus, although VDM formation took place in both liver and skeletal muscle, its destruction was found to be restricted to the liver. In contrast with healthy liver slices maintained under aerobic conditions, normal liver slices previously exposed to anaerobiosis for 2 hours were found to have lost all or most of their capacity to destroy VDM when subsequently restored to an aerobic environment. Thus, the anaerobiosis which led to the formation of VDM by liver had the additional undesirable consequence of causing a concomitant deterioration of its VDM-inactivating mechanism.

These *in vitro* observations led to an investigation of the state of the VDM-inactivating mechanism of liver during both hyper- and hyporeactive phases of hemorrhagic shock. Livers removed during the hyperreactive, reversible stage were found to retain a normal inactivating capacity. Those removed during the hyporeactive, irreversible stage had sustained a marked impairment of their inactivating mechanism. The development of this hepatic defect during the hyporeactive stage would account for the persistence of VDM in the circulation in spite of transfusions which should, at least temporarily, have restored oxidative conditions in the liver.

Other experimental procedures have also served to relate the liver to the *in vivo* inactivation of VDM Blood VDM levels were followed after the intravenous administration of a highly purified beef liver VDM to total or partially eviscerated rabbits.[7] In some of the animals, all the abdominal viscera had been removed; in others, the liver or kidneys, or both, were intact. In the absence of both liver and kidney, the injected VDM persisted in the blood stream; in the presence of the liver alone, the

VDM which was initially present in concentrations comparable to those in shock, disappeared from the circulation within an hour after administration. These results were to be expected from the *in vitro* demonstration that VDM-inactivation was a function of the liver. It was, however, quite unanticipated that VDM should also disappear with equal rapidity from the hepatectomized rabbit with intact kidneys. In view of the invariable failure of kidney slices to inactivate VDM *in vitro,* other mechanisms were investigated which might account for the disappearance of VDM under these circumstances. One was found to be the renal excretion of VDM, considerable amounts appearing in the urine. Thus, the virtual cessation of renal blood flow during drastic hypotension is deleterious for the shocked animal, not only because of the resultant exclusion of renal VEM from the circulation, but also because of the loss, through anuria, of an important mechanism for removing VDM from the blood stream. By the use of the partially eviscerated rabbit, it was possible to observe the unfavorable effects of anoxia on the hepatic VDM-inactivation system *in vivo.* The abdominal viscera were removed with the exception of the liver, whose sole blood supply came from the hepatic artery. When this artery was ligated for 90 minutes and then released, VDM of hepatic origin appeared in the blood stream in increasing concentrations and persisted until *exitus.* Apparently, the VDM-inactivating system had been so damaged by the preceding period of anoxia that it was unable to clear the blood stream of VDM upon restoration of the arterial blood supply to the liver.

An attempt was made to relate the overall respiratory metabolism of the liver to the capacity to destroy VDM. Normal liver slices which had lost this capacity as a result of previous anaerobic incubation were found to have sustained a profound reduction in oxygen consumption, which usually fell to 25-35 per cent of the control values Conversely, livers removed during the hyperreactive phase of shock and which had retained this function, consumed oxygen at a normal rate. This apparent correlation between overall oxidative capacity and inactivation of VDM was, however, not borne out by similar studies with liver removed during the hyporeactive stage. Here, the total loss of inactivating capacity was frequently associated with a reduction in oxygen consumption of as little as 15 to 20 per cent below the average control values The inactivation of VDM would, therefore, appear to be a function of some specific, and probably enzymatic, system in the liver which is extremely sensitive to anoxia.

This assumption is supported by the successful preparation of a cell-free extract from normal liver which has proved to be capable of destroying VDM on aerobic incubation *in vitro.* It effects this destruction under oxidative conditions much more rapidly than an equivalent amount of viable liver tissue and, like normal liver slices, is ineffective under anaerobic conditions. It has been found to consist of a heat-labile apoenzyme which is non-dialyzable, and a heat-stable dialyzable co-enzyme. These two factors may be separated by dialysis and on reconstitution regain their inactivating capacity. The addition of the co-enzyme alone to normal livers which have lost their inactivating capacity through previous anaerobic incubation, serves to restore this capacity. However, the damage to the inactivating mechanism in livers from shocked animals requires for its correction the addition of both the apo- and the co-enzyme.

Inactivation of VEM. Slices of both normal kidney cortex and liver were found capable of inactivating VEM, from whatever source obtained, on aerobic incubation *in vitro,* the inactivation by renal tissue taking place with much greater rapidity. It has already been pointed out that VEM also disappears during prolonged anaerobic incubation of normal kidney tissue, as well as during the virtual anoxia which prevails in the kidney *in vivo* during prolonged periods of drastic hemorrhagic hypotension. It still remains uncertain whether the inactivation of VEM under oxidative conditions and during anaerobiosis are both mediated through the same mechanism. We are inclined to think that the aerobic inactivation of VEM represents a normal function of the healthy kidney, whereas its anaerobic destruction may result from the proteolytic action of by-products of cellular degeneration which is contingent upon prolonged renal anoxia (FIGURE 3).

As with the VDM-inactivating system in liver, previous exposure to anaerobiosis impairs the renal mechanism for the oxidative inactivation of VEM. These observations suggest that the integrity of both hepatic and renal inactivating systems is dependent on the maintenance of an oxidative type of metabolism.

Recapitulation. By the utilization of *in vitro* procedures, more specific information was acquired as to the sites and mode of origin of the humoral vasotropic substances which appeared in the course of hemorrhagic shock. The *in vitro* results were in complete agreement with the data provided by the bioassay of the saline extracts of shocked tissues as regards the sites of origin. VEM formation was limited to the kidney and VDM to liver and skeletal muscle. The only environmental condition which was found to be necessary for the formation of both vasotropic factors was tissue anoxia. The rate of formation of VEM by the kidneys was rapid, whereas VDM formation took place more slowly in liver and skeletal muscle. VDM formation under anaerobic conditions increased progressively with time in both liver and skeletal muscle. Quite another picture was seen with respect to VEM formation by the anaerobic kidney: the peak concentrations were reached in 45-60 minutes, after which VEM progressively fell and was completely absent after 2-3 hours of anaerobic incubation. Mechanisms for the inactivation of VEM and VDM were revealed by *in vitro* studies, which established the existence, in liver, of an enzyme system for the oxidative inactivation of VDM, and in kidney of a mechanism for the oxidative inactivation of

VEM. Liver also possessed to a lesser extent than kidney the capacity to inactivate VEM oxidatively. The oxidative and anaerobic destruction of VEM by kidney is believed to be mediated by different mechanisms, the former being a physiological function, the latter a by-product of tissue degeneration. It was possible to demonstrate an influence of the shock syndrome on the liver-inactivating mechanism for VDM. During the hyperreactive or reversible phase, the liver retained a normal inactivating capacity, while hyporeactivity was invariably accompanied by a significant reduction in this function. It was also possible to demonstrate the removal of VDM from the blood stream by the healthy liver *in vivo* and to prove that a period of hepatic anoxia both *in vivo* and *in vitro,* if sufficiently prolonged, led to the loss of this inactivating capacity. The renal mechanism for inactivating VEM is likewise damaged by prolonged anoxia. In the case of the liver, a cell-free extract consisting of a heat-labile apoenzyme and a heat stable co-enzyme has been prepared and found capable of oxidatively inactivating VDM. In addition to hepatic inactivation, VDM was found to be removed from the blood by excretion into the urine.[7]

DISCUSSION

In recent years, there has been a growing awareness of the inadequacy of the fluid-loss concept as an explanation for many of the phenomena of clinical and experimental shock. These embrace not only the state of irreversibility to fluid replacement therapy seen clinically, and regularly induced by certain standard experimental procedures, but also the decreased tolerance to fluid loss and hypotension in traumatic as compared with hemorrhagic shock, as well as the fatal outcome of tourniquet shock in animals whose limbs have been tightly taped so as to prevent significant fluid loss. As a consequence, attention has been directed at the possible deleterious effects of metabolic derangements or abnormal tissue products resulting from the reduced oxygen tensions prevailing during shock. The search for a possible toxic factor, initiated by Cannon, Bayliss, *et al.*[8] during the first World War, has hitherto been unsuccessful in revealing any agent, whether of tissue or bacterial origin, whose regular appearance and circulatory effects are such as to relate them specifically to the sequence of vascular phenomena characteristic of the conventional shock syndrome. Nor has it been possible to link to these circulatory changes the variety of disturbances in intermediary metabolism which arise in the course of experimental shock.[9]

However, the humoral vasotropic factors recently found by Zweifach, Chambers, and their associates[5] to be regularly associated with the specific patterns of vascular behavior which differentiate the hyper- and hyporeactive stages of shock, would appear to provide a satisfactory explanation for the many phenomena for which the fluid-loss concept has proved inadequate. The effects of these vasotropic factors on the terminal vascular bed are such as to suggest a causal relationship between VEM and the initial compensatory vascular reactions on the one hand, and on the other hand between VDM and the subsequent decompensatory vascular responses to drastic hypotension, which lead to increasing refractoriness to fluid replacement therapy. The experiments which are the subject of the present report served to elucidate the sites, and modes of origin and destruction of these humoral vasotropic principles. They have also provided additional evidence of their causal, as well as temporal, relation to the vascular behavior which differentiates the hyper- and hyporeactive stages of experimental shock.

We are now in a position to synthesize the results of the *in vivo* and *in vitro* experiments into a dynamic concept relating these vasotropic principles to the evolution of the hemorrhagic shock syndrome. In the development of this concept, it will be necessary to include data provided by our studies of vasotropic factors in tourniquet shock[1] as well as the observations of others on changes in blood flow during shock, particularly in the kidney and liver.

The Role of Hepato-Renal Vasotropic Principles in Hemorrhagic Shock. THE HYPERREACTIVE STAGE. Following graded hemorrhage, there promptly ensues a series of vascular compensatory reactions designed to conserve the reduced blood volume for the maintenance of vital structures, such as the heart, lungs, and nervous system, which are immediately essential for survival (TABLE 4). This is accomplished in two ways: by the curtailment of blood flow to the skin, skeletal musculature, and abdominal viscera, particularly the kidney; and by alterations in the functional capacity of the entire peripheral vascular bed. This reduction in blood flow leads to varying degrees of tissue anoxia and to a variety of metabolic derangements[9] whose significance for the shock syndrome still remains obscure except for the vasotropic principles which are elaborated under these conditions. Since these principles originate only in kidney, liver, and skeletal muscle, discussion will be confined to the response of these tissues to the reduced blood flow during the initial phase of hemorrhagic hypotension.

Kidney. The drastic curtailment of renal blood flow which follows hemorrhage has been well established by many studies, particularly those of Van Slyke and his associates,[10] which have revealed the quantitative relation between the degree of hypotension and the extent to which the renal blood flow is reduced. As a result, the oxidative requirements of the kidney are not met, and anaerobic processes are initiated which lead to the rapid formation of VEM and its release into the circulation. By the time the humoral concentration of VEM is measurable by the rat mesoappendix test, characteristic changes are observed in the behavior of the terminal vascular bed. These consist of an enhanced reactivity to epinephrine and an increased vasomotion of the terminal metarterioles and precapillary sphincters. This type of vascular hyperreactivity serves to confine the blood flow to direct thoroughfare channels, thus ensuring the adequate return of the venous blood to the larger veins. As long as the extent of the hemorrhage is such as not to depress the blood pressure below levels of 70 mm. Hg, VEM will persist in the blood stream and maintain this type of vascular hyperreactivity. The causal as well as the temporal relation between humoral VEM and this compensatory type of vascular behavior has been established by the exclusion of the kidney from the circulation just prior to the induction of hemorrhage. Under these conditions, VEM does not appear in the blood and the stage of vascular hyperreactivity fails to develop.

TABLE 4

VASCULAR-METABOLIC DISTURBANCES DURING THE HYPER- AND HYPOREACTIVE STAGES OF IRREVERSIBLE HEMORRHAGIC SHOCK IN ANESTHETIZED DOGS

	Compensatory or hyperreactive	Decompensatory or hyporeactive
Blood pressure level	60 mm. Hg x 120 minutes	40 mm. Hg x 90 minutes
Vascular state	hyperreactive to epinephrine	hyporeactive to epinephrine
Bloodborne principles (rat assay)	VEM—high with progressive reduction	VDM—progressive increase
Liver		
Gross appearance	pale	engorged
Blood flow	reduced but adequate	inadequate for oxidative metabolism
QO_2 *in vitro*	normal range	moderate reduction
VDM content	none	progressive increase
VDM inactivation in O_2	unimpaired	progressive impairment
Kidney		
Blood flow	ischemic, inadequate for oxidative metabolism	none, virtual anaerobiosis
QO_2 *in vitro*	essentially normal	moderate to marked reduction
VEM content	high	progressive reduction
VEM formation *in vitro*	normal	progressive failure
VEM inactivation in O_2	normal	impaired
Skeletal muscle		
Blood flow	moderate ischemia	complete ischemia
Vasotropic principles	mild VEM, bloodborne	VDM—progressive increase, muscle anoxia x time

Liver. The concomitant reduction in hepatic blood flow during this phase of hemorrhagic shock is apparently well tolerated by the liver. Its oxygen consumption remains normal, there is no formation of VDM, and the mechanism for inactivating VDM is unimpaired. These findings indicate that, in contrast to the kidney, the blood flow to the liver, although reduced, is still adequate to maintain an oxidative type of metabolism. The difference in response of the two organs to the reduced blood flow is believed to be attributable to the much higher oxygen requirements of the kidney, and to the fact that the blood supply to the liver is ordinarily much in excess of its oxidative need. Thus, Engel, Harrison, and Long[11] have shown that the hepatic artery is able to provide, alone, for its metabolic requirements. During this stage, the liver is believed to be oxidatively inactivating humoral VEM, but at too slow a rate to reduce the concentration in the blood stream, which is well maintained by continuous renal formation and release into the circulation.

Skeletal muscle. During the same period, the ischemic skeletal musculature is forming VDM. However, the rate of formation is slow, particularly at the low temperatures which usually prevail in the extremities. Furthermore, it is doubtful whether even these small amounts reach the general circulation. If they do, they are either oxidatively destroyed by the liver or masked by overwhelming amounts of humoral VEM. In traumatic shock induced by the application of tourniquets to the lower extremities for 5-10 hours, considerable concentrations of VDM are attained in the occluded extremities and released into the blood stream. This is evident from the analysis of femoral vein blood. However, during the hyperreactive phase, these larger amounts of VDM which enter the blood stream are masked by the even higher concentrations of humoral VEM. Proof for this statement is furnished by experiments with arenal animals in tourniquet shock. Under these conditions, which exclude renal VEM from the circulation, significant amounts of VDM of skeletal origin have appeared in the blood stream within 12 minutes of the release of the tourniquets.

As long as significant amounts of VEM persist in the circulation and the state of vascular hyperreactivity is maintained, the animal remains recoverable by transfusion.

THE TRANSITIONAL PHASE. When more drastic hypotension is now induced in anesthetized animals by additional bleedings sufficient to lower blood pressure levels to below 60 mm. Hg and maintain this reduced pressure, a series of tissue and humoral changes are initiated which result in the abolition of the hyperreactive phase and the progressive development of vascular hyporeactivity. The interval between these two stages may be termed the transitional phase. It is characterized by the return of normal reactivity to the terminal vascular bed. The blood flow is no longer confined to thoroughfare channels but is again allowed to enter the capillary bed through the precapillary sphincters, which resume their normal periodic contraction and relaxation, in contrast to their largely contracted state during the hyperreactive phase. This state of normal vascular reactivity is, however, highly disadvantageous for an animal with a reduced blood volume, since it leads to a further reduction in both blood volume and pressure, by the diversion of blood into the sluggish capillary circulation.

During the transitional phase, the blood gives a neutral reaction by the rat mesoappendix test. This does not necessarily indicate the absence of vasotropic principles. In fact, on many occasions it has been possible, by fractionation procedures, to demonstrate small amounts of both VEM and VDM, whose mutual neutralization has been responsible for the absence of detectible vasotropic effects.[4] These findings suggest that this phase is not always a separate episode but may be the result of a waning hyperreactive and a developing hyporeactive stage.

This transition is attributable to the further reduction in hepatic and renal blood flow that is consequent to the induction of more drastic hypotension. At these lower blood pressure levels, there is a virtual cessation of the renal blood flow, and, as a consequence, renal VEM is locked within the kidney and can no longer enter the blood stream. At the same time, there is a progressive reduction in the concentration of the residual humoral VEM, to which the liver contributes by its oxidative destruction of this vasotropic principle. The blood flow to the *liver* has now decreased to such an extent as to be no longer adequate for oxidative metabolism; anaerobic processes are initiated, and VDM formation begins. Liver biopsies taken at this time, when the blood has become neutral, contain appreciable amounts of VDM, further evidence of the *in vivo* origin of this principle in this organ. Some deterioration of the VDM inactivation system is apparent even at this stage. *Skeletal muscle* contains small but increasing amounts of VDM as a result of the prolongation of the ischemia.

In contrast to the anesthetized animal, the unanesthetized animal in whom the same degree of profound hypotension has been induced by graded hemorrhage, persists in the hyperreactive stage until circulatory failure sets in, and remains recoverable by transfusion. Obviously, since there is no other source of humoral VEM than the kidney, a sufficient blood flow must persist to that organ by which renal VEM is constantly delivered to the general circulation, despite blood pressure levels which abolish renal blood flow in the anesthetized animal. The validity of this inference is supported by the manner in which occlusion of the kidneys alters the shock syndrome in the unanesthetized animal. The hyperreactive stage is abolished and there occurs an early transition to hyporeactivity, and eventual irreversibility, in similar fashion to the behavior of the anesthetized animal in drastic hypotension.

THE HYPOREACTIVE PHASE. With the prolongation of drastic hypotension, the transitional phase is soon superseded by vascular hyporeactivity. This is marked by the appearance of increasing amounts of VDM in the blood stream and by the onset of decompensatory changes in the terminal vascular bed (TABLE 4). These consist in a marked reduction both in the reactivity of the terminal metarterioles and in their spontaneous vasomotion. As a result, peripheral vasoconstriction is further reduced and blood is diverted in increasing amounts into the capillaries and venules, where it remains stagnant. A morbid circulatory cycle is set up by which the reduced blood volume is increasingly lost into the ever-expanding capillary bed.

Liver. In the liver, increasing amounts of VDM appear as a consequence of the progressive hepatic anoxia, which is now sufficient to enforce a virtually anaerobic type of metabolism. Hepatic concentrations of VDM continue to exceed those in the blood and, by virtue of their local action on the hepatic vascular sphincters, lead to the intrahepatic pooling of blood and thereby to a further reduction in effective blood volume. Parallel with this, there ensues a progressive deterioration and eventual failure of the VDM-inactivating mechanism of the liver, as a result of the anaerobic destruction of its enzymatic components (TABLE 1). This deterioration is much more marked than the concomitant reduction in oxygen consumption of the liver, which generally amounts to 15-25 per cent and suggests that the inactivating system is more sensitive to anoxia than the general oxidative mechanisms.

Skeletal muscle. During this stage, there is a further accumulation of VDM in the ischemic skeletal musculature.

The paramount reason for the failure of transfusions during this stage is to be found in their inability to liberate the terminal vascular bed from the decompensatory effects of VDM. *The crucial defect which underlies this failure, and which transfusions cannot reverse, resides in the deterioration of the VDM-inactivation system of the liver.* The mere restoration of oxidative conditions by fluid replacement may temporarily interrupt VDM formation. However, *in vitro* studies have shown that the return of the liver from animals in hyporeactive shock to an aerobic environment for periods up to 3 hours is without benefit to the deteriorated inactivating mechanism, unless there has been supplied, in addition to oxygen, a liver extract containing the specific apo- and coenzyme for the inactivation of VDM. It is also apparent, from *in vitro* studies, that when the damage to the inactivating system is extreme, VDM may continue to form even under aerobic conditions.

Further evidence of the importance of the liver for the development of irreversibility is provided by viviperfusion experiments of Frank, Seligman, and Fine.[12] Dogs subjected to prolonged hemorrhagic hypotension were perfused from the carotid artery of a donor, through the jugular or splenic vein, for a period of 2 to 3 hours. It was possible to prevent the development of irreversibility by viviperfusion of the liver, whereas controls perfused through the jugular vein died soon after transfusion.

Additional factors contributing to irreversibility arise from the damage sustained by the kidney during prolonged anoxia. This damage has two undesirable sequels. The loss of the capacity to form VEM on the restoration of renal blood flow during transfusion deprives the organism of the benefit inherent in the neutralizing effect of this vasotropic factor on VDM. The second consequence of the renal damage is in the suppression of its excretory function. The normally functioning kidney can clear the blood of large amounts of VDM with great rapidity by excreting it into the urine.[7] The anuria of the late stage of shock has, for that reason, more significance for survival than has hitherto been attributed to it.

Relation of VEM to the Crush Syndrome. In this connection, one is tempted to speculate on the possible relation to the crush syndrome of the deleterious effects of prolonged anoxia on the renal mechanism for forming VEM. Van Slyke and his associates have shown that 3 to 4 hours of complete anoxia lead to renal failure and to histological changes comparable to those seen in the kidney in the crush syndrome.[10] *In vitro* studies have established the direct formation of VEM by renal cortical cells, very likely those of the distal convoluted tubules (Unpublished observations). It is a reasonable inference that the local concentration of VEM in the kidney at the site of origin is high and that this vasotropic factor acts on the renal vascular bed just as it does on the peripheral circulatory apparatus. Adequate intraglomerular pressure is essential for normal renal excretory function. It is achieved by constriction of the efferent glomerular arterioles. The observations of Richards[13] indicate a greater response of the efferent than of the afferent glomerular arterioles to epinephrine. The enhancement of epinephrine reactivity by renal VEM should therefore assist in the maintenance of intraglomerular filtration pressures adequate for normal renal excretory function. The damage to the renal capacity to form VEM which follows the prolonged anoxia of the hyporeactive stage of shock, or the renal occlusion procedure of Van Slyke, would exclude this powerful vasotropic principle from participating in the maintenance of adequate intraglomerular pressures. This should lead to inadequate glomerular filtration and constitute an important factor contributing to the renal insufficiency observed under these conditions. It is of interest that the major histological defects in this type of kidney are found in the distal convoluted tubule which our preliminary studies indicate may be a major site of VEM formation.

A recent study of Bywaters[14] points to the involvement of the liver as well in the crush syndrome. A high proportion of livers from cases of crushing injuries or skeletal trauma showed necrosis of the central or mid-zonal cells of the liver lobules which appeared to date from the time of injury. The degree of necrosis was proportional to the duration of hypotension and was attributed to hepatic anoxia.

Vasotropic Factors in Tourniquet *versus* Hemorrhagic Shock. No adequate explanation has hitherto been furnished for the poorer tolerance to equivalent degrees of fluid loss and hypotension exhibited by animals in traumatic as compared with hemorrhagic shock. Our observations would attribute this difference to the greater participation of VDM of skeletal muscle origin in traumatic shock. High concentrations of VDM are attained in the occluded limbs and can be shown to enter the circulation and adversely influence the development of the shock syndrome. The most striking demonstration of the relation of skeletal VDM to traumatic shock is provided by the development of irreversible shock in animals whose limbs have been tightly taped or encased in plaster to prevent any significant local fluid loss after release of the tourniquets. The escape into the circulation of VDM formed by the muscles of the tourniqueted extremities has been shown to be directly responsible for the initiation of the sequence of events which eventually leads to the circulatory collapse of the animal. From these and other experiments cited above, we have been led to conclude that the participation of skeletal, in addition to hepatic, VDM in the evolution of traumatic shock, is the major factor which differentiates this syndrome from hemorrhagic shock.

Therapeutic Implications. It is apparent from these experimental results that the correction of a state of shock which has become irreversible to the present type of fluid replacement therapy requires the utilization of other measures in addition to transfusion. These measures, in order to be effective, should correct the derangements which have developed in the hepatic vasodepressor and renal vasoexcitor mechanisms. The immediate therapeutic goal consists in the liberation of the vascular bed from the decompensatory influences of VDM; the ultimate goal, in the arrest and reversal of the progressive dysfunction of the liver with respect to VDM formation and inactivation, which results from prolonged anoxia. The first might be achieved by the inactivation of humoral VDM by appropriate agents such as the cell-free liver extract which has proved capable of inactivating VDM under aerobic conditions *in vitro*. The attainment of the second may be dependent upon the administration of substrates favorable to the regeneration of the liver inactivation system. Another approach would consist in supplying VEM in amounts sufficient to insure the resumption of the compensatory type of vascular reactivity, which in turn would lead to the restoration of an adequate blood flow to the tissues, particularly to liver and kidney.

Objection has frequently been expressed to the use of the term, irreversible shock, in view of the more favorable results which are being obtained by the use of more massive and frequent whole blood transfusions. However, the term, irreversibility, appears to be a legitimate description of the present status of replacement therapy for certain well-defined states of experimental shock and will have served its purpose if it has focused attention on those factors which are responsible for its development. It may well be that the more favorable results recently observed to follow the extensive use of massive transfusions have been achieved by permitting the survival of the patient long enough for a regeneration of the hepatic and renal vasotropic mechanisms to take place. The benefit derived from whole blood transfusions may, therefore, consist not only of the maintenance of aerobic conditions in the organism, but also in the supply of essential precursors of these enzymatic systems.

The General Significance of VEM and VDM. There are two possible interpretations of the role of VDM in shock. It might be regarded as an abnormal, "toxic" by-product of hepatic anoxia, or as a physiological principle whose toxic effect results from its excessive concentrations in shock. The latter represents our present viewpoint.

This is supported by the existence in the liver, of specific enzymatic mechanisms both for its production and destruction. The same considerations would apply to the participation of renal VEM in the shock syndrome. As in the liver, there exist in the kidney enzymatic mechanisms not only for the formation but for the inactivation of this vasotropic principle. The converse action of VEM and VDM on the terminal vascular bed has brought us to the hypothesis that these vasoexcitor and vasodepressor principles are oppositely-acting components of a homeostatic mechanism participating in the regulation of peripheral blood flow and blood pressure. The vascular arrangements in both kidney and liver are such as to favor the development of the intermittent ischemia necessary for the production of these vasotropic principles. This, in addition to the presence of systems in both organs for their formation and inactivation, provides an ideal chemico-physiological basis for such a homeostatic mechanism. In the course of experimental shock, the magnitude and abruptness of the circulatory disturbances lead to the initial preponderance of the renal vasoexcitor principle and the eventual predominance of the hepatic vasodepressor. These circumstances represent drastic disturbances of equilibrium. Under more normal circumstances, circulatory homeostasis would undoubtedly be accomplished by more subtle alterations in the equilibrium between these vasotropic principles.

BIBLIOGRAPHY

1. **Shorr, E., B. W. Zweifach, & R. F. Furchgott**
1945. Science **102**: 489.
2. **Zweifach, B. W., R. E. Lee, C. Hyman, & R. Chambers**
1944. Ann. Surg. **120**: 232.
3. **Zweifach, B. W., B. E. Lowenstein, & R. Chambers**
1944. Am. J. Physiol. **142**: 80.
4. **Zweifach, B. W., R. G. Abell, R. Chambers, & G. H. A. Clowes**
1945. Surg. Gynec. Obst. **80**: 593.
5. **Chambers, R., B. W. Zweifach, B. E. Lowenstein, & R. E. Lee**
1944. Proc. Soc. Exp. Biol. & Med. **56**: 127.
6. **Wiggers, C. J., & J. M. Werle**
1942. Proc. Soc. Exp. Biol. Med. **49**: 604.
7. **Baez, S., B. W. Zweifach, A. Mazur, & E. Shorr**
1947. Proc. Soc. Exp. Biol. & Med. **64**: 154.
8. **Cannon, W. B.**
1935. Traumatic Shock. D. Appleton & Co. New York, N. Y
9. **Wilhelmi, A. E., & C. N. H. Long**
1948. Ann. N. Y. Acad. Sci. **49** (4): 605.
10. **Van Slyke, D. D.**
1948. Ann. N. Y. Acad. Sci. **49** (4): 593.
11. **Engel, F. L., H. C. Harrison, & C. N. H. Long**
1944. J. Exp. Med. **79**: 9.
12. **Frank, H. A., A. Seligman, & J. Fine**
1946. J. Clin. Invest. **25**: 22.
13. **Richards, A. N.**
1929. Methods and Results of Direct Investigations of the Function of the Kidney. Williams & Wilkins Co. Baltimore.
14. **Bywaters, E. G. L.**
1946. Clin. Sci. **6**: 19.

EFFECTS OF HEMORRHAGE ON THE KIDNEY

By Donald D. Van Slyke
The Hospital of The Rockefeller Institute for Medical Research, New York, N. Y.

THE EFFECT OF HEMORRHAGE ON THE COMPLETENESS OF THE EXTRACTION OF PARA-AMINOHIPPURIC ACID (PAH) BY THE KIDNEY AND ON THE ACCURACY OF THE PAH CLEARANCE AS A MEASURE OF RENAL BLOOD FLOW

Dr. Homer Smith and his collaborators have shown that the plasma clearance of para-aminohippurate serves as an approximate measure of the volume of plasma that flows through the kidneys per minute. For the PAH clearance to equal exactly the renal plasma flow, it would be necessary for the kidneys to extract, from the plasma of the blood, 100 per cent of the PAH, so that the plasma of the renal venous blood would be entirely free of hippurate. In order to ascertain how accurately the hippurate clearance measures the renal plasma flow both under normal

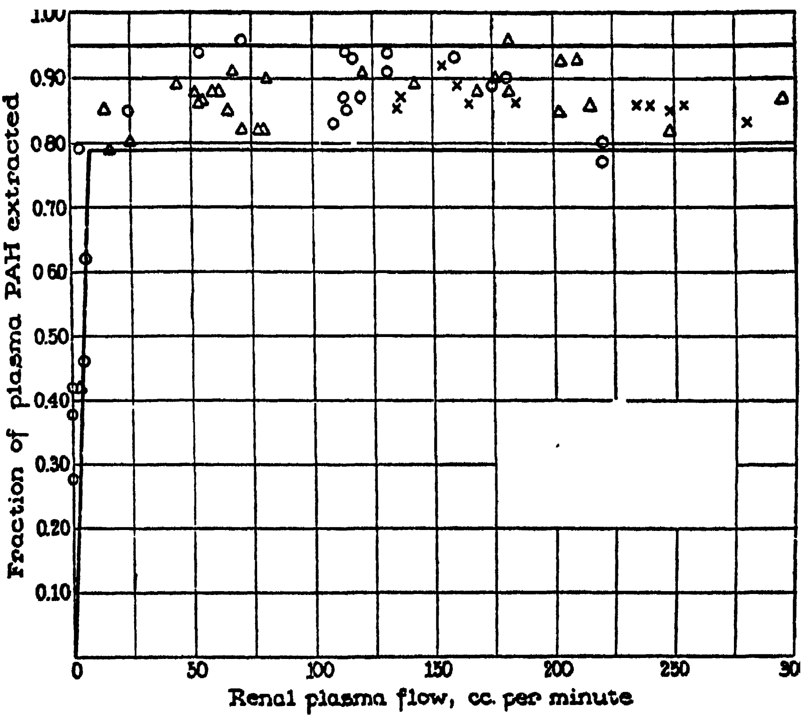


Figure 1.

conditions and after hemorrhage, it was necessary to measure the completeness of the extraction. This was accomplished[1] by comparing the PAH concentration of plasma from the renal vein with the PAH concentration of the arterial plasma. It was found, as shown in FIGURE 1, that an average of 87 per cent of the plasma PAH was extracted by the kidneys, and that when the renal plasma flow was decreased by severe hemorrhage, this 87 per cent extraction was completely maintained until the flow of blood through the kidney fell below 3 per cent of normal. FIGURE 2 shows renal plasma flows calculated as 1.15 times the hippurate clearance compared with actual plasma flows. The actual plasma flows in liters per minute were measured by dividing the amount of hippurate excreted in one minute by the amount extracted from one liter of plasma. It is seen, from FIGURE 2, that the plasma flow, even when extremely decreased by hemorrhage, can be estimated as 1.15 times the hippurate clearance.*

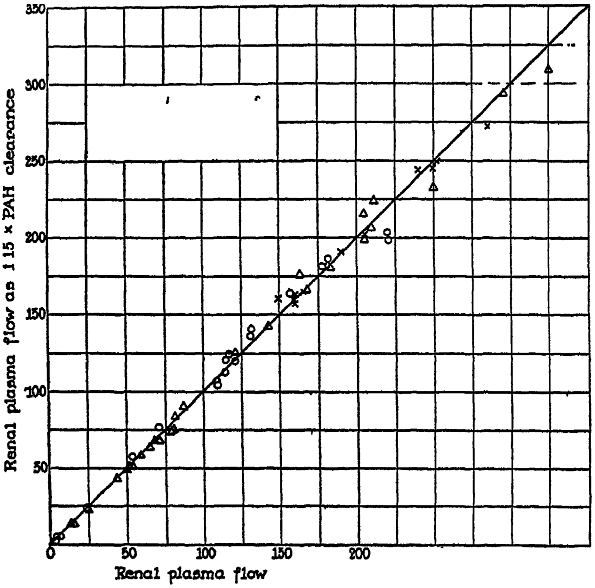


FIGURE 2.

The greater part of the hippurate excreted by the kidney is excreted by the tubules, so that the completeness of extraction of the hippurate from the plasma serves as a measure of the functional activity of the tubular cells. It appears that this activity is maintained even when, as the result of hemorrhage, the renal blood flow is decreased to 3 per cent of normal. The phenomenon could be explained by assuming that when the renal blood flow is thus decreased, it is completely shut off from part of the nephrons, while those which continue to be perfused function in an entirely normal way. This assumption would also explain the fact, observed in our studies, that a great reduction in renal blood flow during hemorrhagic shock is not accompanied by a significant decrease in the oxygen saturation of the renal venous blood. In the body as a whole, when the circulation is decreased, the tissues compensate for the retarded blood flow by extracting a larger part of the oxygen from the blood, so that the venous blood becomes dark and shows an increased degree of oxygen unsaturation. In the kidneys, on the contrary, the blood flow can be reduced to a very low level and still the renal venous blood maintains its normal, nearly arterial degree of redness and its oxygen content.[2] This also could be explained by the diminished volume of blood passing through an equally diminished number of nephrons at a normal rate with the rest of the kidney receiving relatively little blood. The constant and nearly complete extraction of PAH indicates that no increased fraction of renal blood by-passed the nephrons in the manner recently described by Trueta.[3]

---

* The above conclusion applies only when the duration and severity of the hemorrhagic shock and the resultant renal ischemia are not so great that the kidneys suffer organic damage of the type that can cause uremia after recovery from shock. When such damage occurs, the tubules are so injured that the kidneys extract much less than the usual 87 per cent of hippurate from the plasma, and 1.15 times the hippurate clearance is much less than the renal plasma flow (for example, see FIGURE 10). The experiments reported in FIGURES 1 to 9 represent results of acute hemorrhagic shock, not of sufficient severity (except perhaps in FIGURE 4) to cause such injury.

THE IMMEDIATE EFFECT OF HEMORRHAGE IN CAUSING RENAL ISCHEMIA AND THE SUBSEQUENT RESTORATION OF RENAL BLOOD FLOW BY PERIPHERAL CONSTRICTION

In FIGURE 3 are shown the effects of repeated hemorrhage on the arterial blood pressure and on the para-aminohippurate clearance of a dog. It is evident that severe and rapid hemorrhage causes an immediate reduction in the hippurate clearance and therefore the renal plasma flow, but that there quickly follows a rise both in arterial blood pressure and in the renal blood flow. In FIGURE 3, this phenomenon is demonstrated in three successive hemorrhages during the course of three hours in a single dog. After a fourth hemorrhage, restoration of the renal blood flow (para-aminohippurate clearance) was incomplete, but infusion of a volume of plasma equal to only one-fourth of the blood that had been withdrawn restored the hippurate clearance to practically normal.

The apparent explanation of these phenomena is that, immediately after a sudden hemorrhage, the kidney shares with the rest of the body a decrease in blood flow, but that within a certain period, which may be only a few minutes, constriction of the peripheral circulation sets in, and the circulation of the kidney is more or less completely restored. Presumably this peripheral constriction is caused by the vasopressor substance studied by Shorr, Zweifach, Chambers, and their collaborators. Evidence that the total vascular bed is constricted after severe hemorrhage, is the fact, shown in FIGURE 3, that after hemorrhage continued to the verge of death, replacement by a plasma infusion of only one-fourth of the blood withdrawn entirely restored the circulation of the kidney.

In FIGURES 3 and 4, PAH represents para-aminohippurate clearance, B. P. arterial blood pressure, and Cr the creatinine clearance.

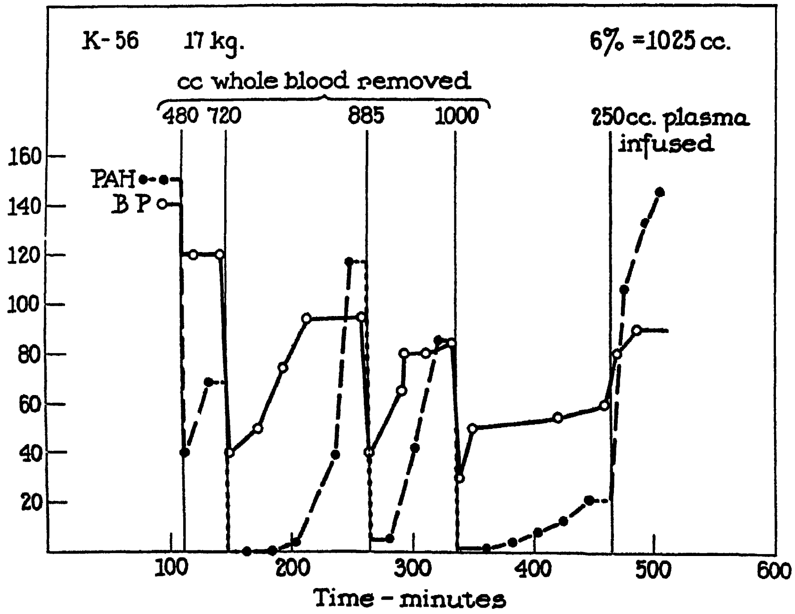


FIGURE 3.

FIGURE 4 shows the results of a hemorrhage so severe that, during a period of several hours, blood flow through the kidney was not sufficient to be measured. Thereafter, replacing all the blood lost resulted in only a slow and incomplete restoration of renal blood flow and of glomerular filtration as measured by the creatinine clearance. The partial restoration of renal function was only temporary, followed by a rapid decline both of renal function and of blood pressure. It is apparent that this animal had reached the stage of advanced shock in which peripheral constriction gives way to dilatation and the shock becomes irreversible, a phenomenon which is explainable by the replacement of the vasoconstrictor material by vasodepressor material noted by Shorr, Zweifach, and Chambers.

EFFECTS OF GRADUAL BLOOD LOSS ON RENAL BLOOD FLOW AND GLOMERULAR FILTRATION: EVIDENCE OF COMPENSATORY CONSTRICTION OF THE EFFERENT VESSELS OF THE NEPHRONS

FIGURE 5 presents data which permit an analysis of the effects of hemorrhage on renal function. Blood was withdrawn from the dog in successive portions, as indicated by the curve at the bottom of the figure. The values for plasma flow were determined by dividing the amount of para-aminohippurate excreted in the urine per minute by the amount removed from one liter of plasma during passage of the kidney, as measured by simultaneous analyses of the plasma of arterial and renal venous blood. Creatinine clearances were determined and interpreted as measures of the filtration rate in terms of volume of glomerular filtrate formed per minute. Dividing the filtration rate by the plasma flow gives the fraction of water in the plasma that is filtered during passage of the glomerulus. This fraction is normally about 25 per cent of the plasma water. The curve for the filtered fraction is indicated in the figure. It is evident that renal plasma flow undergoes a progressive fall after withdrawal of the first 25 cc. of blood per kgm. The arterial blood pressure, however, is maintained until the gradual withdrawal is more than 40 cc. per kgm. It is evident that during this period, peripheral constriction has served to keep up the general arterial blood pressure, but that the constriction, to some extent, includes the kidney as well as peripheral tissues. A striking compensatory effect is the marked increase in the percentage of plasma water filtered in the glomerulus. This rises from 26 per cent to 38 per cent (FIGURE 5) during the first 4 hours of the experiment. After this period, there is a rapid fall in both

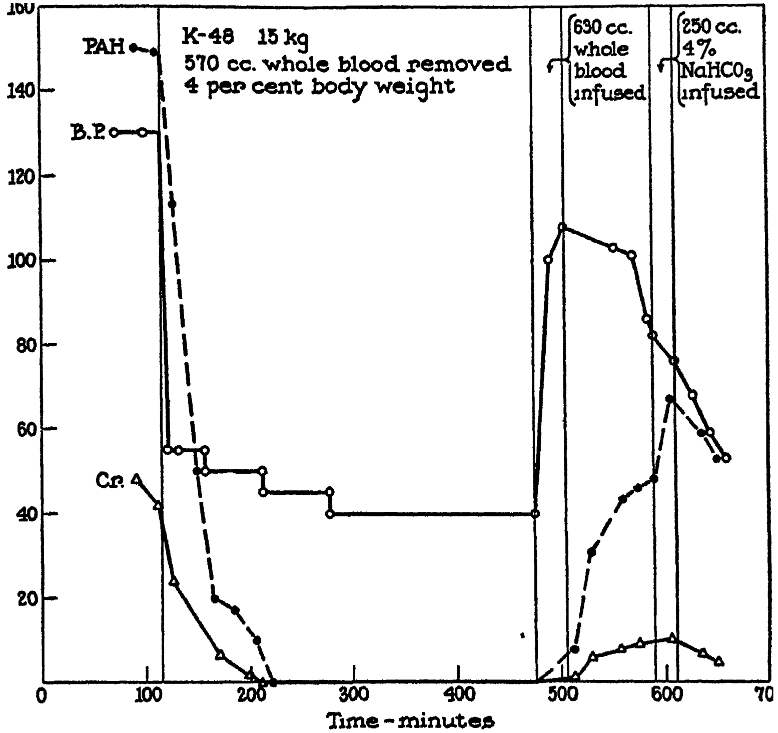


FIGURE 1.

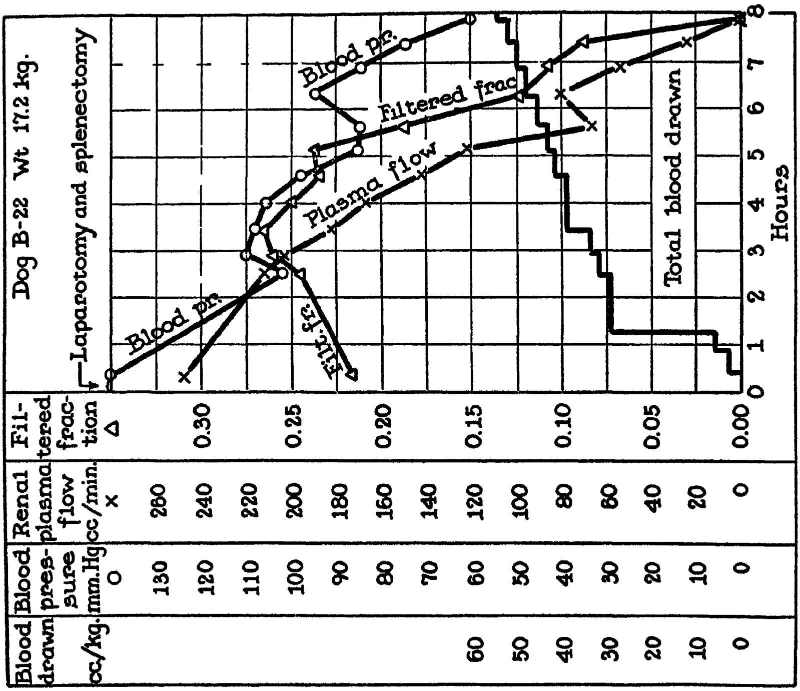


FIGURE 6.

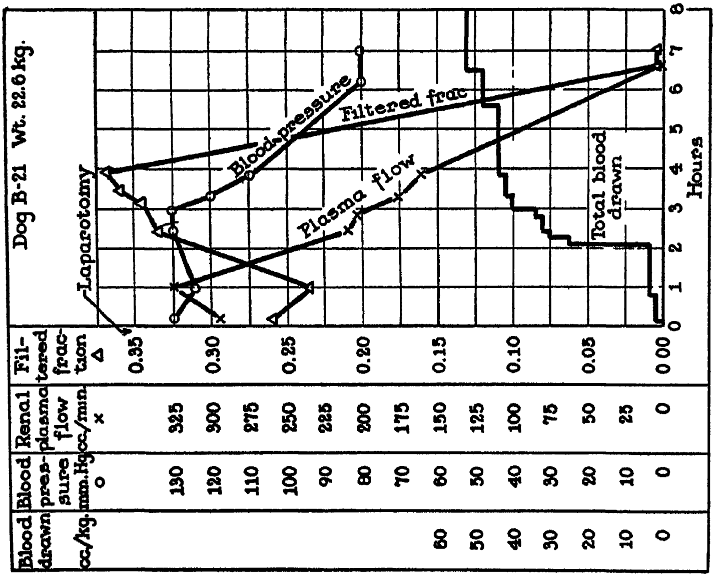


FIGURE 5

renal plasma flow and in the filtered fraction, so that the excretory function of the kidney falls to almost zero. This debacle of the kidney occurred despite the fact that the blood pressure was still maintained at 80 mm. It appears that, when hemorrhage is sufficiently severe to demand a maximal peripheral constriction in order to maintain blood pressure and blood flow through the heart and brain, the blood flow to the kidney is shut off, as is the flow to the skeletal muscles, and that the function of the kidney is suppressed in order to maintain the vital central organs.

In FIGURE 6, another example of similar phenomena is shown. During the period from the 5th to the 11th hour of the experiment, the blood pressure, despite continuous withdrawal of blood, is raised from 80 to 90 mm. The filtered fraction, which has fallen to 13 per cent, is raised to 25, presumably because constriction of the efferent vessels of the kidney is more marked than that of the afferent. Eventually, after the 7th hour, the shut-down of the kidney occurs. FIGURE 7 is another example of the same phenomena.

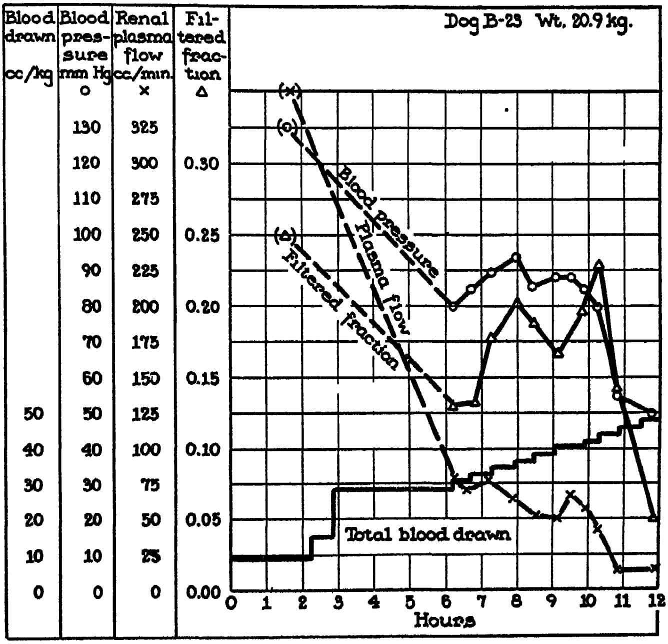


FIGURE 7.

RELATIVE EFFECTS OF HEMORRHAGE ON THE FUNCTION OF THE KIDNEY AND THE LIVER IN THE DOG

The question whether death is more likely to be caused by uremia or by failure in other vital organs, may be due to the relative extent to which renal function, compared with that of other organs, is affected. FIGURE 8 presents data, unpublished so far, which give some comparison of the relative effects of hemorrhage on the kidney and on the liver. In the normal dog the function of the liver in oxidizing uric acid is so active that the plasma uric acid level is normally maintained below 0.5 mg. per 100 cc. The plasma ammonia concentration is normally kept below 0.1 mg. per 100 cc., also by action of the liver, according to available data. In FIGURE 8, it is seen that a hemorrhage sufficient to reduce renal plasma flow to zero for 2 hours was not accompanied by significant rises in either the ammonia or uric acid content of the plasma. However, when a second portion of blood was subsequently withdrawn, lowering the arterial blood pressure to 25 mm., the plasma ammonia nitrogen and uric acid began to rise rapidly. This rise continued until, after 90 minutes, the blood withdrawn was replaced. This replacement was followed by an immediate drop of the uric acid and ammonia contents of the plasma, accompanied by a restoration also of the renal blood flow. It appears from such experiments that in the dog, the function of the kidney is abolished by hemorrhage before that of the liver.

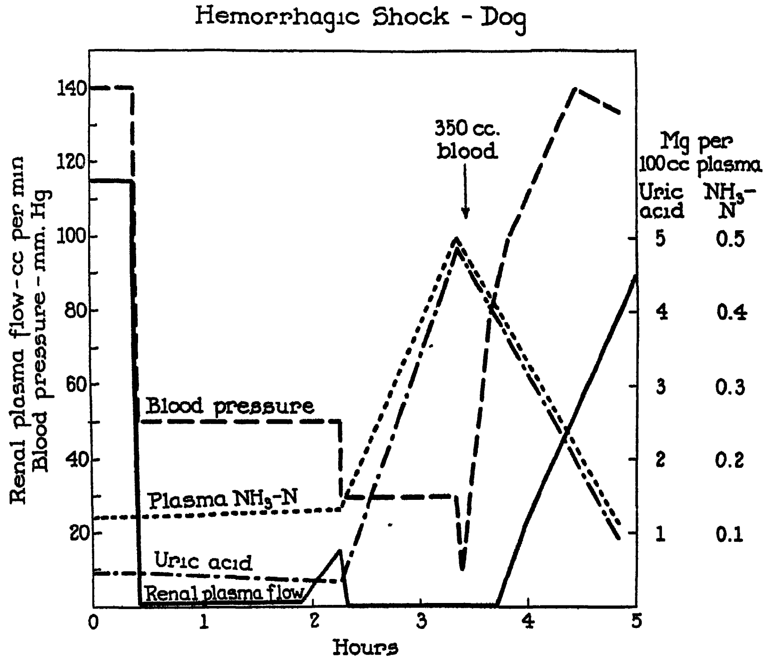


FIGURE 8.

FIGURE 9 gives the plasma amino acid and urea concentration values obtained in the same experiment. It will be seen that the amino acid nitrogen after the second withdrawal of blood shows the same rapid rise that was demonstrated by ammonia and uric acid. Since removal of amino acids from the circulation has been shown to be a function of the liver, the behavior of the amino acid curve strengthens the supposition that the liver function did not fail greatly until after the second blood withdrawal. Work of many investigators, especially Man and his collaborators, has shown that urea is formed exclusively in the liver, and that when the kidney function is discontinued, blood urea will rise steadily as a result, as long as the liver continues to function. During the first 2 hours, the blood urea does, in fact, rise steadily, indicating that the liver was functioning. Thereafter the blood urea curve becomes relatively flat, again indicating approximate cessation of liver function.

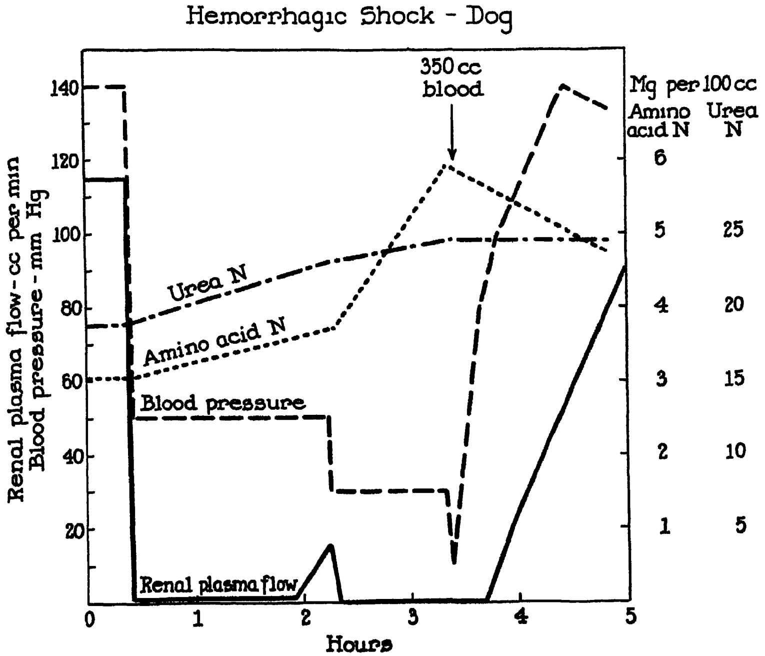


FIGURE 9.

RENAL FUNCTION AFTER THE RELEASE OF TEMPORARY CLAMPING OF THE RENAL ARTERY

The relative orders, with regard to time, in which the kidneys and the liver cease to function during progressive hemorrhage, does not necessarily indicate the order in which damage to these organs from ischemia becomes irreversible. The kidneys of the dog have been shown by Phillips and Hamilton[4] to endure complete occlusion of the renal artery for as long as 3 hours without irreversible damage. After 4 hours, however, the damage is practically always irreversible, and the animal dies a few days later in uremia, without return of renal function. FIGURE 10 shows the slow recovery of renal function as measured by the hippurate clearance, and the renal plasma flow after 2 hours' clamping of the renal artery. The effect of clamping the renal artery is more severe than the effect of hemorrhage sufficient to reduce the renal blood flow below a measurable rate for the same length of time. It is possible to depress dog renal blood flow to nearly zero by hemorrhage for several hours, and then to obtain almost immediate return of function by restoration of the lost blood. It is probable that during the depression caused by hemorrhage, some small trickle of blood continues to get to the kidney, enough to prevent damage to the nephrons as rapid as that which is obtained by complete closure of the renal artery.

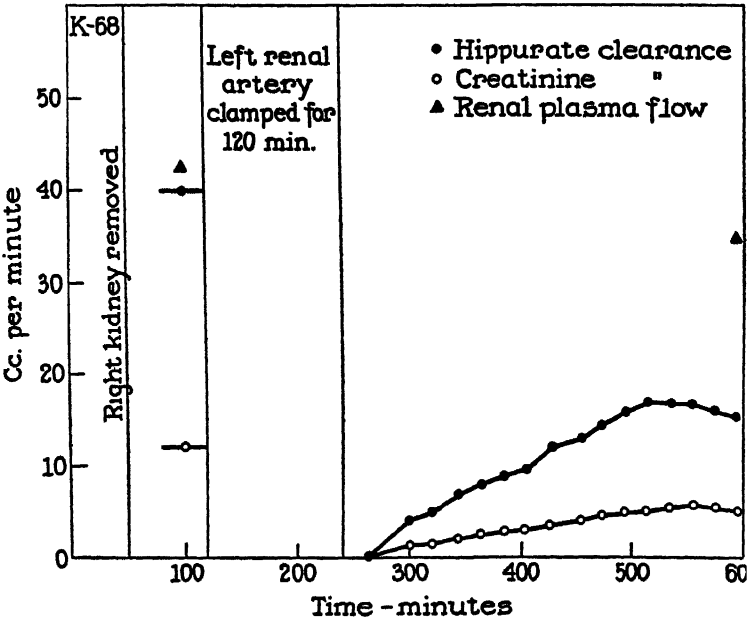


FIGURE 10.

In a given animal species, the question whether hemorrhage can produce uremic death appears to depend on the relative ability of the kidney, compared with other vital organs, to withstand the effects. In man, death in uremia a week or more afterwards has been observed to follow shock caused by hemorrhage, trauma, or dehydration such as occurs in Asiatic cholera. In these conditions, anuria, caused by the decreased blood volume, fails to be relieved with restoration of the lost blood or fluid, or, if some flow of urine is reestablished, it is of such low volume and concentration that it does not suffice to prevent uremia. The condition is similar to that observed in dogs in which the renal artery was closed for more than 3 hours.

In dogs, it has not been possible to produce uremia by hemorrhage.

The apparent reason is that the kidney is more resistant, relative to the other vital organs, in the dog than it is in man, so that if hemorrhage is made sufficiently severe and prolonged in the dog to cause irreversible renal failure, the animal dies from other damage, usually circulatory, without surviving long enough to develop uremia.

SUMMARY

After sudden hemorrhage of 20 to 30 cc. per kg., the sequence of events observed in dogs under nembutal anesthesia was a drop in arterial blood pressure to 50-60 mm., with cessation of measurable renal blood flow and excretion, followed quickly by partial or nearly complete restoration of central blood pressure and of renal function. The restoration appeared to be attributable to constriction of the extrarenal peripheral vessels, and to indicate that renal circulation was favored at the expense of extrarenal peripheral circulation.

However, if further progressive hemorrhage followed during a period of several hours, over an approximate range of 30 to 40 cc. of total blood loss per kg., this blood loss was accompanied by progressive decrease in renal blood flow, although arterial blood pressure might be maintained above 100 mm. During this period, the fraction of plasma water filtered in the glomeruli, as measured by the creatinine extraction, tended to increase, and thereby to uphold the volume flow of glomerular filtrate despite the shrinkage in renal blood flow. The circulatory phenomena during this period apparently included partial constriction of the afferent renal vessels, adding renal to peripheral constriction in the endeavor to maintain central blood pressure, while the efferent renal vessels constricted still more than the afferent ones, to produce the compensatory increase in the fraction of plasma water filtered.

When hemorrhage surpassed a certain limit, about 40-45 cc. per kg., or when blood lost by somewhat smaller hemorrhage was not replaced for some hours, or when muscle trauma exceeded a certain limit, both renal blood flow and the fraction of plasma water filtered fell to almost zero levels. This renal debacle might occur when the central blood pressure was still 80-100 mm. It appeared that, by maximum afferent renal constriction, the organism at this stage temporarily strangled the kidneys in an effort to maintain central blood pressure. If not too severe or prolonged, the condition in this stage was still reversible by infusion of blood or plasma, or by spontaneous recovery.

The relations between central blood pressure, renal blood flow, and glomerular filtration in shock caused by muscle trauma were similar to the relations noted in shock caused by hemorrhage.

The results reported were obtained by Phillips, Dole, Hamilton, Emerson, Archibald, and the author[1, 2, 4] in investigations at the Hospital of The Rockefeller Institute carried out under the sponsorship of the Institute, the Navy, and the Committee of Medical Research of the O.S.R.D.

BIBLIOGRAPHY

1. **Phillips, R. A., V. P. Dole, P. B. Hamilton, K. Emerson, Jr., R. M. Archibald, & D. D. Van Slyke**
   1946. Effects of acute hemorrhagic shock on renal function of dogs. Am. J. Physiol. **145:** 314.
2. **Dole, V. P., K. Emerson, Jr., R. A. Phillips, P. B. Hamilton, & D. D. Van Slyke**
   1946. The renal extraction of oxygen in experimental shock. Am. J. Physiol. **145:** 327.
3. **Trueta, J., A. E. Barclay, K. G. Franklin, P. Daliel, & M. M. L. Prichard**
   1946. Renal pathology in the light of recent neurovascular studies. Lancet **237:** 251.
4. **Phillips, R. A., & P. B. Hamilton**
   1946. Duration of renal ischemia in dogs required to produce damage of lethal degree. Proc. Fed. Am. Soc. Exp. Biol. **5** (2): 80.

METABOLIC CHANGES ASSOCIATED WITH HEMORRHAGE

By Alfred E. Wilhelmi and C. N. H. Long

Department of Physiological Chemistry, Yale University, New Haven, Connecticut

The study of the dynamics of the circulatory system necessarily takes first place in the physiology of hemorrhage. It has led to the development of therapeutic measures which, when applied in good time, can bring about recovery in many instances after severe hemorrhage associated with profound shock. The study of metabolism after hemorrhage has as its principal objective the solution of problems raised by the limitations of the present therapeutic methods. This solution has to be based on an understanding of the metabolic changes accompanying, and perhaps responsible for, the so-called "irreversible" state of shock. One such change, the production of vasodepressor material by the anoxic liver, was described and discussed in one of the preceding papers.[1] An important part of the metabolic studies consists in discovering to what extent and how rapidly the defects brought about by severe hemorrhage can be repaired, since, if these changes can be reversed, a hopeful basis for developing supplementary therapeutic measures is established.

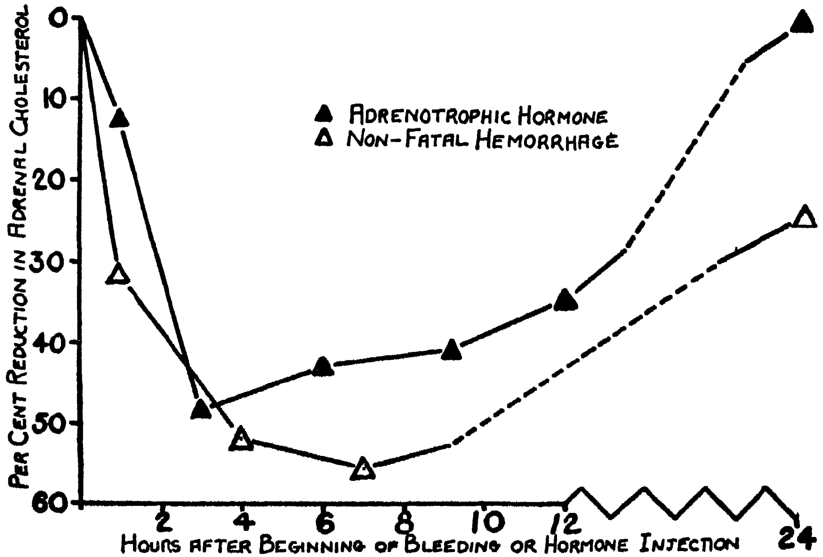


FIGURE 1. Depletion of adrenal cholesterol following a non-fatal hemorrhage, after administration of a single dose of adrenocorticotrophic hormone (cf.[8]).

The studies of the metabolic changes associated with hemorrhage that have been carried out in our own and other laboratories now make it possible to differentiate between two patterns of change in animals after bleeding. The first group of changes is characteristic of the response of animals to stresses in general, in which the principal condition to be met is an increased demand for energy which may have to be sustained for a long time. The response of the rat to a non-fatal hemorrhage (loss of about 1 ml. of blood per 100 cm.^{2} of body surface in the course of an hour) provides a good illustration of these characteristic metabolic changes.[2]

Shortly after bleeding has begun, there is a marked fall in the concentrations of cholesterol and ascorbic acid in the adrenal cortex (FIGURES 1 and 2). This change is reversed within a few hours, and by 24 hours the concentrations of both substances have returned to or near the normal level. A similar response is evoked by the injection of pituitary adrenocorticotrophic hormone. In the hypophysectomized animal, a non-fatal hemorrhage does not bring about any significant changes in adrenal cholesterol and ascorbic acid (FIGURE 3). Hemorrhage, like many other types of stress, leads to stimulation of the pituitary gland and release of adrenocorticotrophic hormone, which in turn

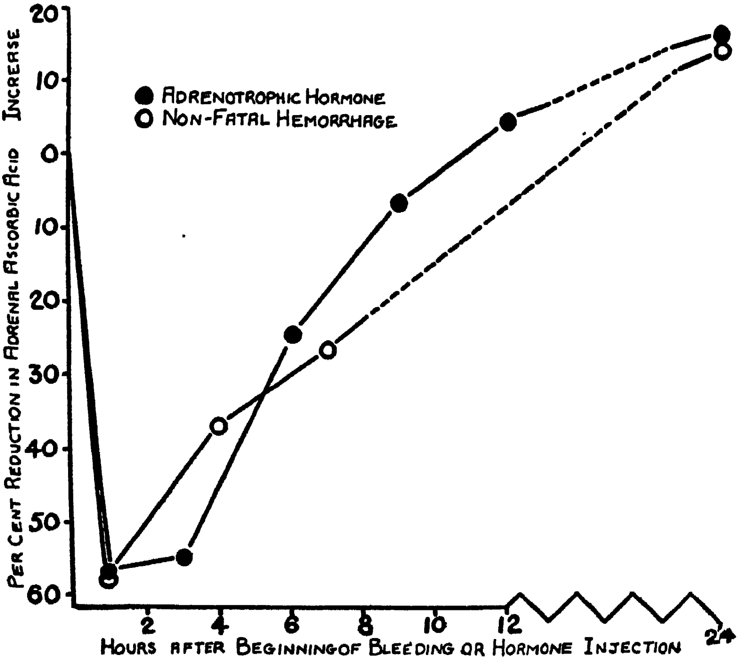


FIGURE 2. Depletion of adrenal ascorbic acid following a non-fatal hemorrhage, after administration of a single dose of adrenocorticotrophic hormone (cf.[3]).

brings about an increase in the activity of the adrenal cortex. The effects of this increased output of cortical hormones cannot yet be assessed in terms of the specific metabolic reactions in which they are involved, but its importance is clearly indicated by the fact that both hypophysectomized and adrenalectomized animals are abnormally sensitive to stresses of all kinds. In general terms, the increased secretion of cortical hormones may be related to an increase in nitrogen catabolism and to increased formation of new carbohydrate.

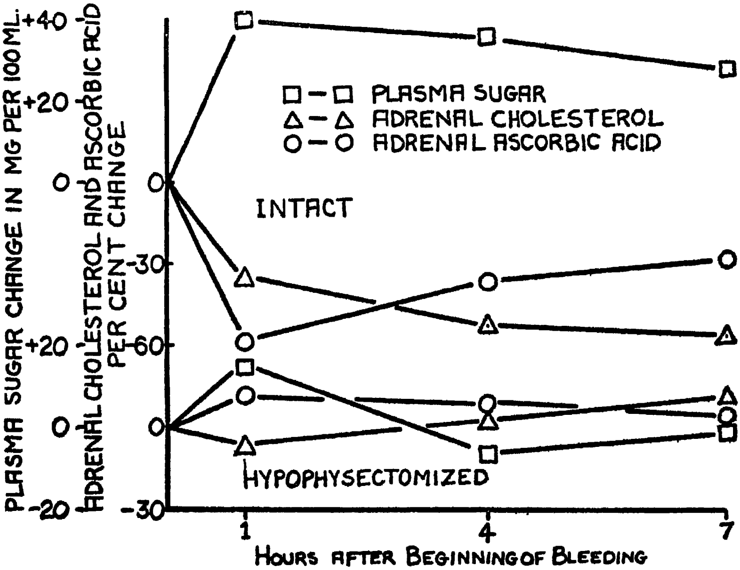


FIGURE 3. Changes in plasma sugar, adrenal cholesterol, and adrenal ascorbic acid of intact and hypophysectomized rats, following a non-fatal hemorrhage (cf.[3]).

The most striking change in the blood after non-fatal hemorrhage is an increase in blood sugar which may be sustained for some hours. There are also increases in blood lactate and pyruvate. However, the lactate/pyruvate ratio is within normal limits, indicating that the metabolic processes in the peripheral tissues are mainly oxidative. These changes are probably due to the reflex secretion of epinephrine, since they are not seen in the adrenodemedullated animal. There is a small decrease in plasma cholesterol, a slight increase in ascorbic acid, and no change in plasma amino nitrogen. Daniel Kline[3] has recently shown that in the dog, after non-fatal hemorrhage, there is a sustained increase in the output of amino acids from the muscles, but that since there is an increased intake of amino acids by the liver, no net change in plasma amino nitrogen takes place.

Additional evidence of an increased protein metabolism after non-fatal hemorrhage has been obtained in the nephrectomized rat.[4] In this preparation, the intravenous injection or infusion of an amino acid mixture is followed by an increased rate of accumulation of urea in the blood (FIGURE 4). The increment is largest in the first hour and subsides to the basal level by the end of the third hour after injection. After a non-fatal hemorrhage, the rate of urea production is sustained at the initial high level for a longer time.* These observations are consistent with the well-known fact that there is an increased nitrogen excretion after recovery from many kinds of injury or stress, including hemorrhage severe enough to have caused shock and temporary anuria.

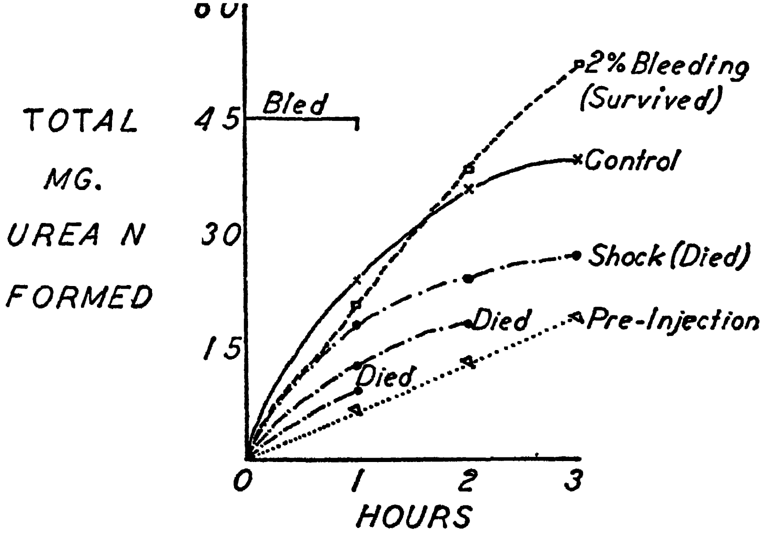


FIGURE 4. The effects of a single intravenous injection of 1.5 ml. of 10 per cent casein hydrolysate (16.5 mg. N) on urea formation in nephrectomized rats, normal, after a non-fatal hemorrhage, and in a fatal case. The pre-injection curve represents the normal hourly rate of urea production prior to injection of amino acids (cf.[4]).

The liver also appears to be normal in other respects. There is no significant change in the concentration of electrolytes and of water.[5] Also, the rates of oxygen uptake, deamination, and urea synthesis by the liver tissue *in vitro* are relatively normal.[6, 7] There is a fall in liver glycogen which is related to the increase in blood sugar, and, since the increased blood sugar is sustained for many hours even in the fasted rat, it is probable that there is new formation of carbohydrate from lactate and amino acids brought to the liver from the peripheral tissues.

These responses of the animal to a hemorrhage from which it can recover spontaneously are similar in kind and in degree to the metabolic changes accompanying severe exercise or other kinds of stress. One of their characteristics is that they seem to be designed to meet an increased demand for energy in circumstances in which the supply of oxygen or of substrate, or both, is relatively and temporarily insufficient. Another important characteristic is that the responses are reversible, so that within a relatively short time the organism returns to its initial steady state.

---

* It may be noted here that Engel has since discovered, and Russell in our laboratory has confirmed the observation, that the high rate of urea production observed in these experiments is not seen if the amino acid solution is carefully sterilized. However, the exaggerated urea production brought about by the injection of non-sterile amino acid solution helps to emphasize the point that, after a non-fatal hemorrhage, the capacity of the liver to synthesize urea from administered amino acids is in no way impaired.

The second group of metabolic changes associated with hemorrhage accompanies the development of shock. It arises as a consequence of the diminished supply of oxygen and nutrients to the tissues. Some of these changes are an exaggeration of the normal metabolic responses to stress. Others are new, and are related to the failing metabolic activities of one or more organs.

In shock, the concentrations of cholesterol and ascorbic acid in the adrenal cortex fall to still lower levels than those attained during the normal response to stress, and they show no tendency to return toward the initial levels (TABLES 1 and 2). This may be a sign of continued intense stimulation of the adrenal cortex, such that the demand for cholesterol and ascorbic acid outruns the supply. Since the rat is capable

TABLE 1

ADRENAL, LIVER, BRAIN, AND PLASMA CHOLESTEROL FOLLOWING A FATAL HEMORRHAGE

Values expressed as mg. per 100 mg. of fresh tissue and mg. per 100 ml. of plasma; each value represents the average and standard error of 5 to 7 animals.

	Controls (not bled)	Bled	
		Intermediate stage	Terminal
Adrenal	5.84 ±0.314	2.58 ±0.414	1.53 ±0.140
Liver	0.391±0.0239	0.325±0.0093	0.330±0.0101
Brain	1.89 ±0.031	1.85 ±0.040	1.93 ±0.041
Plasma	55±2.6	43±1.8	42±3.5

TABLE 2

ADRENAL, LIVER, BRAIN, AND PLASMA ASCORBIC ACID FOLLOWING A FATAL HEMORRHAGE

Values expressed as mg. per 100 gm. of fresh tissue and mg. per 100 ml. of plasma; each value represents the average and standard error of 5 to 7 animals.

	Controls (not bled)	Bled	
		Intermediate stage	Terminal
Adrenal	403 ±15.2	231 ±11.6	205 ±7.9
Liver	24.0 ± 1.10	20.5 ± 2.18	13.2±0.99
Brain	39.6 ± 0.72	41.3 ± 0.69	38.9±1.37
Plasma	0.51± 0.058	2.57 ± 0.128	3.0±0.476

of synthesizing ascorbic acid, this animal is unsuitable for determining whether the supply of the vitamin can be a limiting factor in the response to severe stress. It has recently been observed by Christianne Dosne de Pasqualini[5] that the pretreatment of guinea pigs with large doses of ascorbic acid greatly increases the proportion of animals surviving after a severe, otherwise usually fatal, hemorrhage. It has not yet been demonstrated, however, that the administration of ascorbic acid at a time when the circulation is already seriously diminished can contribute materially to recovery from shock, nor is it certain yet that the functional activity of the adrenal cortex remains unimpaired in severe shock. At present, most of the questions raised by the observation that the adrenal cortex is intensely activated in severe stress and that both cholesterol and ascorbic acid play an important part in this activity, are unanswered.

As shock develops after a fatal hemorrhage, the blood sugar, which may at first increase, begins to fall, owing to depletion of liver glycogen and possibly also to failure of the liver to make new carbohydrate, but especially owing to a greatly increased rate of utilization of carbohydrate by the peripheral tissues.[9] The last effect may be seen clearly in the eviscerated rat (FIGURE 5). In this preparation, the blood sugar usually falls slowly, reaching hypoglycemic levels in 1½ to 2 hours. After hemorrhage, there is a precipitate fall in blood sugar, and the animal dies in convulsions within a short time. At the same time, there is a

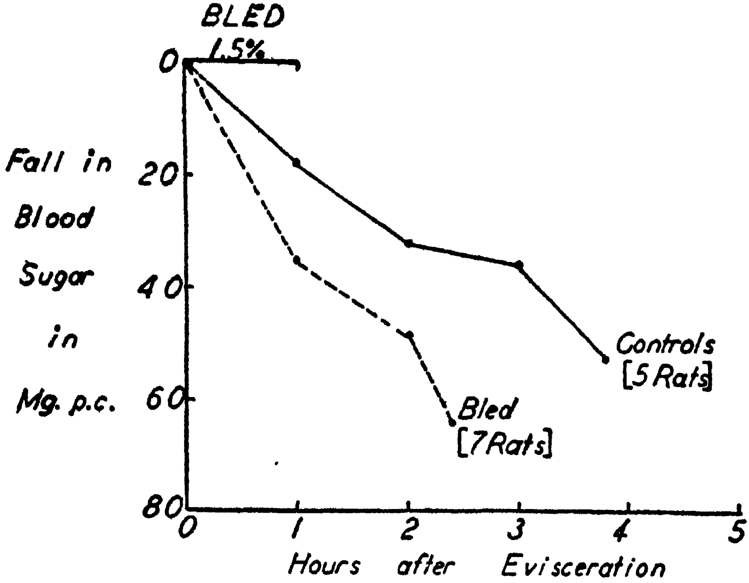


FIGURE 5. The effect of hemorrhage on the blood sugar levels of eviscerate adrenodemedullated rats. The rats bled lost 1.5 per cent of their body weight of blood during the first hour after evisceration (cf.[9]).

large increase in blood lactate, and a moderate increase in pyruvate (FIGURE 6). The high lactate/pyruvate ratio, which is also seen in the

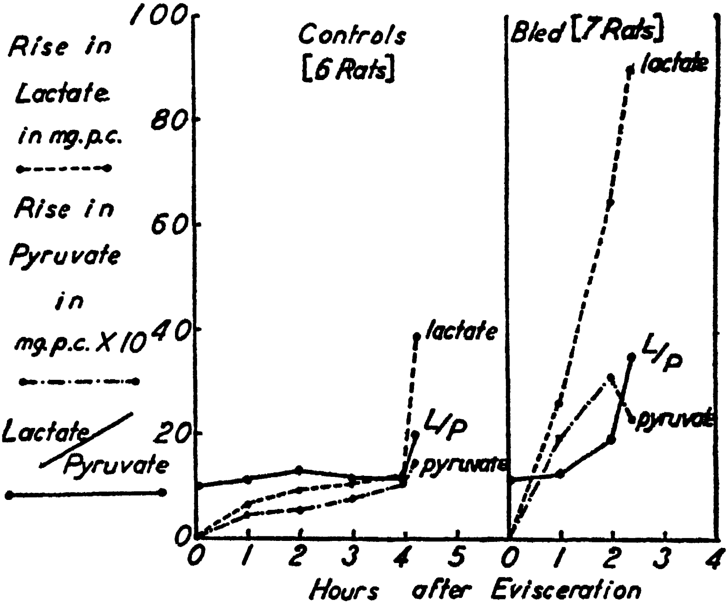


FIGURE 6. The effect of evisceration, and of evisceration plus hemorrhage, on the blood lactate and pyruvate levels of adrenodemedullated rats. The rats bled lost 1.5 per cent of the body weight of blood immediately after evisceration. Convulsions immediately preceded death (cf.[9]).

intact animal in shock,[10] marks a shift toward anaerobic carbohydrate metabolism in the anoxic peripheral tissues. This may be correlated with the increased carbohydrate requirement.

The lactacidemia observed in shock can be associated with the marked fall in carbon dioxide combining power of the plasma. Acidosis is one of the metabolic changes first associated with shock. Cannon and other workers, more than 25 years ago, gave bicarbonate to patients and animals in shock, in the expectation that relieving the acidosis might increase their chances of recovery. During the late war, the effects of alkalinizing agents were studied experimentally by a number of investigators.[11-14] The observations of Wiggers and Ingraham, in a recent paper,[14] are fairly representative of the results obtained. In severe hemorrhagic shock in dogs, the administration of sodium lactate was without beneficial effect, but the administration of sodium bicarbonate resulted in a significant increase in the rate of survival. In almost every instance, bicarbonate brought about a marked increase in the carbon dioxide combining power of the plasma, while lactate did not. The lactacidemia of shock is a sign that lactate is already being produced at a rate faster than it can be removed. In these circumstances, sodium lactate is unlikely to be an effective alkalinizing agent. It may be emphasized that the acidosis of shock is a metabolic acidosis, and that it is associated with a state of insufficient oxygen supply to the tissues. Nothing is gained by administering a substance that must be rapidly metabolized to be effective in relieving the acidosis.

A new and important chemical change in the blood after severe hemorrhage is a progressive rise in amino nitrogen with increasing severity of shock (FIGURES 7 and 11). This change appears at a time when the blood pressure has fallen and remains below about 80 mm. of Hg. It is coincident with a marked fall in portal venous oxygen tension The increase in amino nitrogen can be associated, in part, with a failure, due to diminished oxygen supply, of the capacity of the liver to assimilate amino acids, and in part with an accelerated output of amino acids by the peripheral tissues. In rats in which the gastrointestinal tract is removed and the blood supply to the liver is provided solely by the hepatic artery, the blood amino nitrogen remains at normal levels (FIGURE 8). When the blood supply to the liver is cut off by clamping the hepatic artery, the blood amino nitrogen increases, since there is a slow, steady output of amino nitrogen from the peripheral tissues[15, 16]. If the clamp is released after a short time, the liver can arrest this increase in blood amino nitrogen, but if the period of anoxia is sufficiently long, the liver no longer can remove amino acids from the blood when blood flow through the organ is restored by releasing the hepatic artery. In the intact animal, the major part of the blood and oxygen supply to the liver is provided by the portal vein, and in shock, the rise in blood amino nitrogen can be correlated with the fall in blood pressure and with a decline of portal venous oxygen tension to very low levels, that is to say, with progressive anoxia of the liver tissue (FIGURE 9). The increased

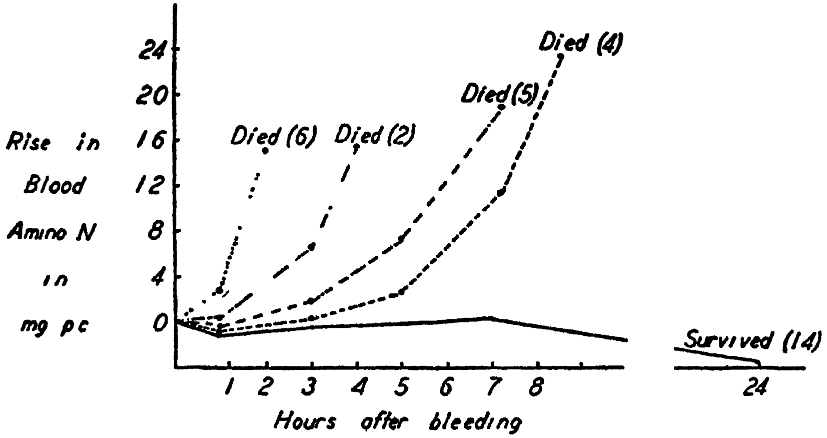


FIGURE 7. The effect of blood loss equivalent to 3 per cent of the body weight in one hour on the blood amino nitrogen levels in normal rats (cf [10]).

peripheral contribution to the blood amino nitrogen in shock may be seen in the eviscerated rat[9]. In this preparation, the slow, steady output on amino nitrogen from the peripheral tissues is greatly accelerated after hemorrhage (FIGURE 10). Kline[8] has recently reported similar changes in plasma alpha-amino nitrogen after fatal hemorrhage in dogs. There is an increase in femoral vein amino nitrogen over the arterial level which persists in spite of a steadily rising level of arterial amino nitrogen. The arterial-portal difference in amino nitrogen increases as well, indicating that amino nitrogen is added to the blood as it passes through the gastrointestinal tract. Although the portal-hepatic differ-

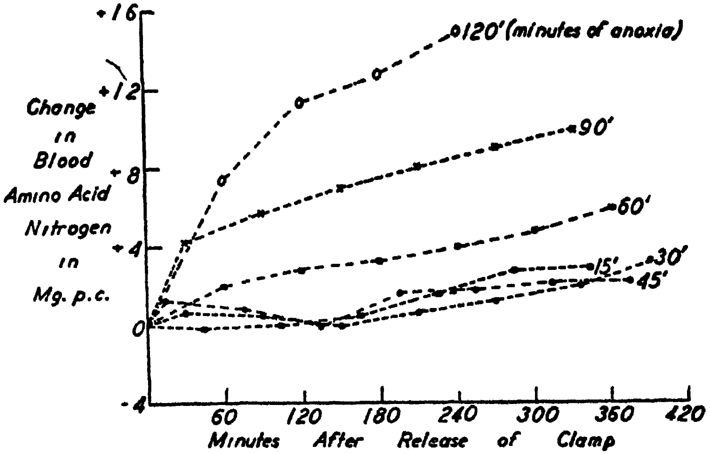


FIGURE 8. The effect of complete occlusion of the hepatic circulation for 15 to 120 minutes on the blood amino nitrogen level of the rat (cf.[10]).

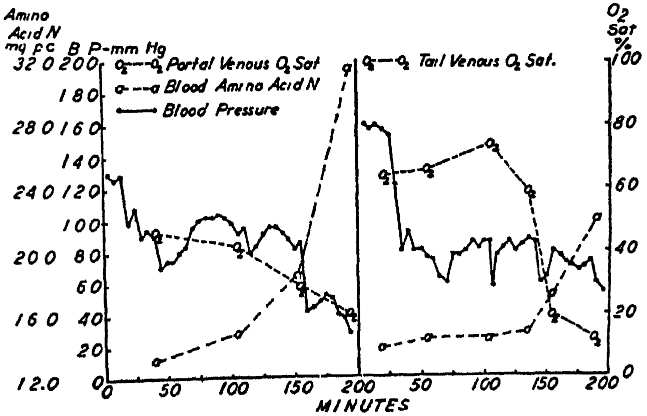


FIGURE 9. The blood amino nitrogen, blood pressure, and oxygen saturations of portal and tail veins in two rats subjected to fatal hemorrhage (cf.[10]).

ences show that the liver increases its intake of amino nitrogen from the blood, the rate of increase is not adequate to prevent a general rise in plasma amino nitrogen.

The rise in blood amino nitrogen is one of the most reliable indices of the severity of shock in the rat[17, 19]. Owing to changes in the hematocrit during shock and to the high concentration of amino nitrogen in the red cells, the change in plasma amino nitrogen is a more dependable measure than the change in whole blood. In the rat, an increase in plasma amino nitrogen of more than 1 mg. per cent by the end of an hour after the onset of bleeding indicates that the animal will die within another hour unless some therapeutic measures are taken (TABLE 3).

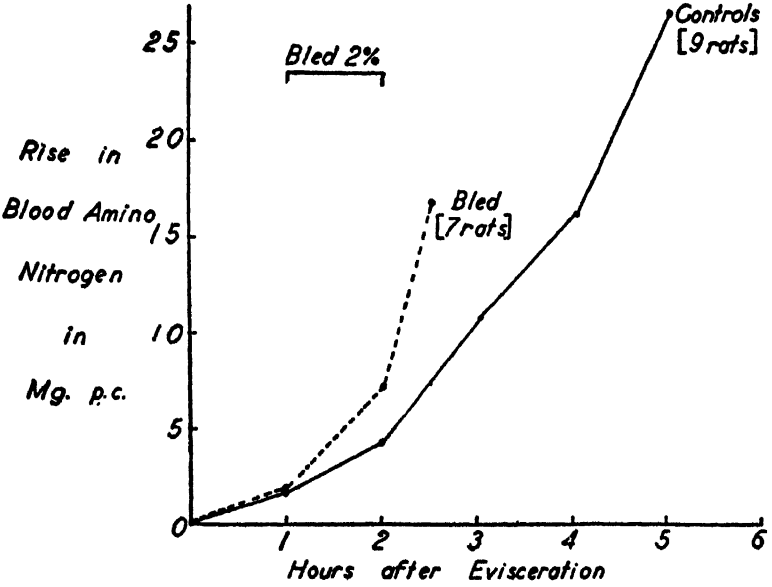


FIGURE 10. The effect of hemorrhage on the blood amino nitrogen of eviscerate rats. The rats bled were subjected to a hemorrhage of 2 per cent of the body weight of blood between the first and second hours after evisceration (cf. [9]).

Such a rise in plasma amino nitrogen occurs when shock is already well advanced, but it may be seen before any "irreversible" change has taken place. The level of the plasma amino nitrogen may also be useful in estimating the degree of recovery from shock. This possibility, however, requires experimental investigation.

Additional evidence that the liver is early and seriously damaged in shock, has been obtained from studies of urea synthesis *in vivo* and *in vitro*,[4, 6] from studies of the rates of deamination and of oxygen intake of the liver tissue,[6] and from determinations of the composition of the liver with respect to water and electrolytes and amino nitrogen.[5, 18] In the nephrectomized rat, subjected to a fatal hemorrhage, the injection or constant intravenous infusion of an amino acid mixture leads to a much lower rate of accumulation of urea in the blood than that found in control animals (FIGURE 3). At the same time, the blood amino nitrogen increases, whereas it usually is unchanged in the controls or even in nephrectomized rats subjected to a non-fatal hemorrhage (FIGURE 11). These observations are supported by the findings that urea synthesis, deamination, and oxygen intake are diminished in the liver tissue of rats in hemorrhagic shock (FIGURE 12). In the intact animal, urea synthesis and deamination seem to be affected in the same degree, since there is an increase in liver tissue amino nitrogen in shock, without an increase in ammonia at a time when urea formation is greatly diminished.[6,18] The disorganization of liver metabolism in shock has

TABLE 3

RELATION BETWEEN THE CHANGE IN PLASMA AMINO NITROGEN DURING BLEEDING AND THE SURVIVAL OF RATS

Change in plasma amino nitrogen* mg. %	Number of rats	Died	Survived	Deaths per cent	Survival time minutes†
—1.0 to —0.1	9	3	6	33	55 to 140 (Avg. 84)
0 to +0.5	20	11	9	55	5 to 72 (Avg. 36)
+0.5 to +1.0	19	14	5	74	6 to 65 (Avg. 33)
+1.0 to +5.6	26	26	0	100	2 to 120 (Avg. 35)
	74	54	20	73	

All animals bled 2.5-2.7 ml. per 100 g. of body weight in 1 hour.

* Change in plasma amino nitrogen from beginning to end of hemorrhage.

† Survival time of animals that died only are included. Time in minutes after end of bleeding.

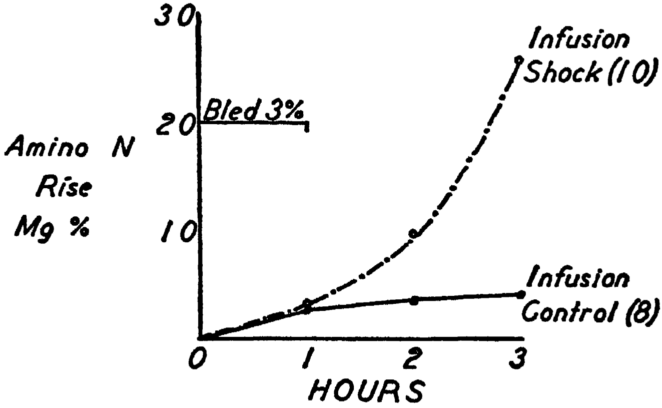


FIGURE 11. The effect on the blood amino nitrogen of continuous intravenous infusion of a 10 per cent casein hydrolysate solution, in control and bled nephrectomized rats. Amino nitrogen plotted as rise (in mg. per cent) over the pre-injection levels (cf.[4]).

also been associated, in the studies of Govier and Grieg,[20-24] with the destruction in the anoxic tissue of the coenzymes, co-carboxylase, cozymase, and alloxazine-adenine dinucleotide, and with the inactivation of some of the specific proteins with which these co-factors operate. A similar inactivation of the enzyme system responsible for the destruction of the vasodilator material has been reported in a previous paper by Shorr.[1]

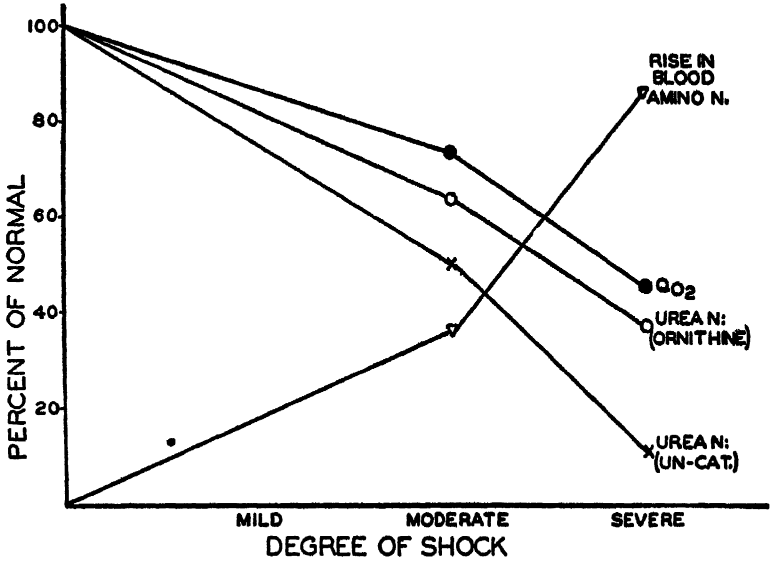


FIGURE 12. Percentage changes with increasing severity of shock after hemorrhage (normal = 100 per cent) in blood amino nitrogen, rate of oxygen uptake of liver slices in phosphate buffer with 0.2 per cent glucose, and urea nitrogen synthesized from ammonia and lactate, and ammonia, lactate and ornithine, by liver slices in bicarbonate buffer (cf.[6]).

All of the evidence indicates that the liver is particularly vulnerable in shock, and that it may suffer damage well before other organs have been seriously affected. One reason for this lies in the nature of the hepatic blood supply, which is unequally divided between the hepatic artery and the portal vein, so that the greater part —60 to 75 per cent cent—of the blood and oxygen supply to the liver is provided by the portal vein. It has been shown, by Blalock and Levy,[25] that a hemorrhage sufficient to reduce the blood pressure only to 80 mm. of Hg may reduce the blood flow through the liver by more than 50 per cent. Wiggers, Opdyke, and Johnson[26] have shown recently that after acute hemorrhage in dogs there is an increase in hepatic resistance relative to mesenteric resistance to blood flow, so that the total blood flow through the liver is likely to be reduced much more than the fall in arterial pressure would indicate. This factor, combined with an early fall in portal venous oxygen tension, must lead to hepatic anoxia relatively soon after a severe hemorrhage.

Other factors contributing to hepatic damage may derive from the normal metabolic activities of the liver, and from its normal response to stress. The liver is an organ of carbohydrate storage and synthesis. One of its functions is to deliver glucose to the bloodstream on demand, and it does this in response to hemorrhage. The liver may, therefore, be depleted of glycogen at a time when its requirement for this substrate is increased owing to the reduction of its oxygen supply. The liver is also an important site of fat oxidation, and judging from the respiratory quotients found by many investigators in liver slices from both fed and fasted animals, fat usually comprises a large proportion of the metabolic mixture being oxidized in the liver. In anoxia, there may be an accumulation of free fatty acids in the liver cells in concentrations high enough to inhibit tissue respiration, since a supply of energy for glyceride synthesis or for maintaining the fatty acids in the phosphorylated form required, according to Lehninger,[27] for their oxidation, may be cut off. This speculation is intended to emphasize the point that there is as yet little detailed evidence illustrating the exact effects of anoxia on tissue respiration.

In the terminal stages of hemorrhagic shock in the rat, the final collapse of the animal is heralded by a well-defined series of respiratory changes which suggest that there might be a serious disturbance in the central nervous system.[28] Our studies on the metabolism of brain tissue slices,[29] as well as the observations of Drabkin and his colleagues[30] on minced brain tissue from animals in shock, yielded no evidence that the brain as a whole had been subjected to any considerable degree of anoxia before the terminal collapse took place. The methods are, of course, crude, and they cannot detect changes in small regions of the brain. The speed of the final collapse, taken together with the fact that an animal will survive if the transfusion is large enough[28] suggests that there is no cumulative defect in brain metabolism due to continuing relative anoxia during shock. It may simply be that one or two vital centers of the brain need only be deprived of their blood and oxygen supply for a relatively brief period. The problem of immediate survival is, therefore, still one of maintaining the blood pressure and oxygen supply adequate, in order to prevent these terminal changes in the central nervous system.

There is, by now, general agreement that the metabolic changes associated with severe hemorrhage are the result of prolonged and progressive tissue anoxia. One of the valuable results of the metabolic studies may be to emphasize the point that even after recovery from shock, a considerable amount of repair must take place, particularly in the liver and the kidney, before the organism can resume the normal course of general recovery. Van Slyke and his colleagues[31] have, in fact, shown that after several renal anoxia, there may not be time for adequate tissue repair to take place before the patient dies of uremia. The liver presents an even more serious problem, since some degree of recovery must be brought about if the production of vasodilator material is to be stopped and its removal facilitated. The importance of maintaining the liver in good condition is illustrated by the ingenious cross-circulation experiments of Fine and his co-workers.[32] Hypotension is induced in dogs by bleeding them out into a pressure bottle, and is maintained to the point at which returning all of the blood lost does not bring about survival. If the liver is perfused with blood from a donor dog during the period of hypotension, all the animals survive. A similar perfusion through the peripheral vessels is without effect on the survival rate. These experiments in prevention of liver damage, after hemorrhage, illustrate how important it may be to bring about recovery of the liver as soon as possible.

For this reason, we have undertaken some studies on the ability of the liver to resist, and to recover from, anoxia.[33, 34] The preparation used was the partially eviscerated rat, with the hepatic blood supply provided only by the hepatic artery. The rate of oxygen intake of slices of the liver tissue was used as an indicator of tissue activity. The effects of anoxia were determined by clamping the artery and then releasing it for a two-hour period (FIGURE 13). It was found that the hepatic arterial blood supply was capable of maintaining normal rates of respiration in the liver. This is in agreement with Engel's observation[16] that this

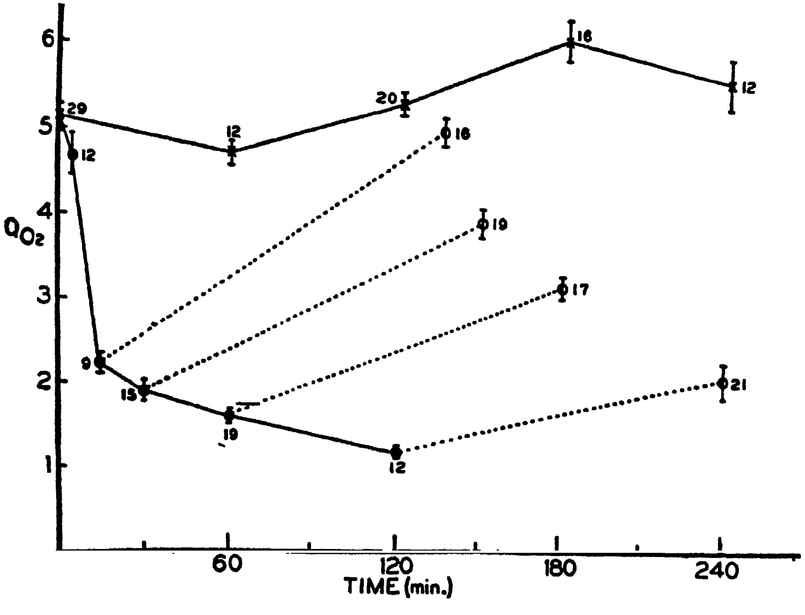


FIGURE 13. Effect of hepatic anoxia *in vivo* on the respiration of rat liver slices *in vitro*. Crosses: control observations on liver slices from operated animals. Solid circles: samples from livers with hepatic artery clamped for 0 to 120 minutes. Hollow circles: samples taken two hours after restoring circulation to the liver following 15, 30, 60, or 120 minutes of hepatic anoxia. Small figures are number of observations. Vertical bars equal standard error of the mean (cf.[33]).

preparation maintains a normal level of blood amino acids. In fasted rats, the liver respiration is sharply diminished after only short periods (15 to 30 minutes) of anoxia. The degree of recovery in 2 hours is inversely proportional to the duration of anoxia, and may be appreciable even after two hours of anoxia. That recovery of respiration may be accompanied by recovery in at least one other respect, may be seen from the experiment, described before,[16] on the effects of hepatic anoxia on the ability of the liver to assimilate amino acids (FIGURE 14). If the changes in blood amino nitrogen, after circulation through the liver is restored, are compared with the increases in blood amino nitrogen that

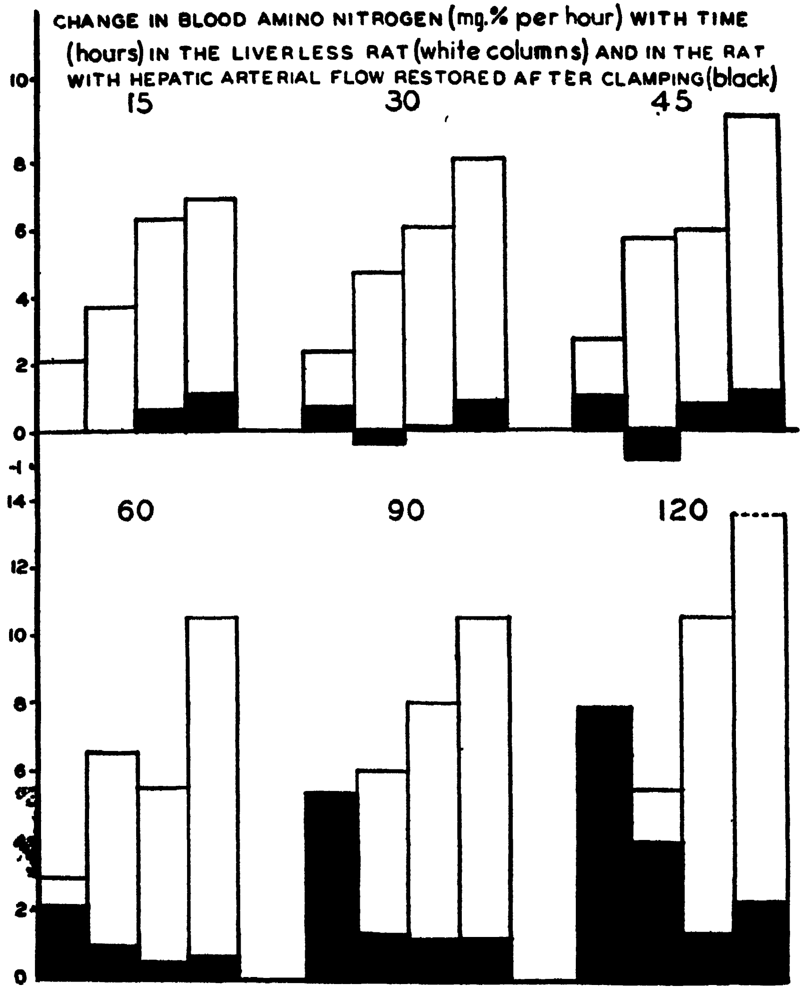


FIGURE 14. Rate of increase of blood amino nitrogen following periods of hepatic anoxia up 120 minutes in length (black columns), compared with the increases observed in the corresponding times in the liverless rat. (Calculated from data presented in [16].)

occur in the liverless animal at the same time, it is seen that even after two hours of anoxia the liver recovers, to some extent, its ability to remove amino acids from the blood.

The liver *in situ,* therefore, appears to have a large native capacity for repair when its normal blood and oxygen supply are restored after a long period of anoxia. A natural question to ask, in this connection, is whether, if adequate blood flow and oxygen supply could be restored in animals in profound shock, the liver might not within reasonable time recover its capacity to destroy the vasodilator material that seems to play so critical a part in the terminal circulatory collapse of the animal. An intimation that this might indeed occur may be obtained from some observations of Sayers, Sayers and Long on the resuscitation of rats in the terminal stages of prolonged hemorrhagic shock.[28] In animals acclimatized at 84° F., transfusions of whole blood equal to 5 per cent of the body weight brought about survival in all of 7 instances, and transfusions of 0.9 per cent sodium chloride solution equal to 10 per cent of the body weight brought about survival in 4 out of 5 instances. Although more work is required before any certain conclusions can be drawn from these experiments, they suggest that, if the volume of circulating blood be made large enough to maintain blood pressure and blood flow at levels adequate to forestall irreversible changes in the central nervous system, enough time might be gained to allow the liver to recover from the effects of anoxia, and to set about reversing what at present seems to be an irreversible situation.

To the students of circulatory dynamics, it may be a source of satisfaction that the students of metabolism should appeal to them for rescue from the metabolic consequences of hemorrhage. It is, however, a sign of the interdependence of these studies, and a token that, as more information is gained on both sides in future experiments, a coherent and inclusive understanding of the physiology of hemorrhage will be attained.

BIBLIOGRAPHY

1. **Schorr, E., B. W. Zweifach, & R. F. Furchgott**
   1945. Science. **102**: 489.
   1948. Ann. N. Y. Acad. Sci. 49 (4): 571.
2. **Sayers, G., M. A. Sayers, T. Y. Liang, & C. N. H. Long**
   1945. Endocrinology **37**: 96.
3. **Kline, D. L.**
   1946. Am. J. Physiol. **146**: 654.
4. **Engel, F. L., & M. G. Engel**
   1946. Ibid. **147**: 165.
5. **Darrow, D. C., & F. L. Engel**
   1945. Ibid. **145**: 32.
6. **Wilhelmi, A. E., J. A. Russell, M. G. Engel, & C. N. H. Long**
   1945. Ibid. **144**: 674.
7. **Russell, J. A., C. N. H. Long, & A. E. Wilhelmi**
   1944. J. Exp. Med. **79**: 23.

8. **de Pasqualini, C. Dosne**
1946. Am. J. Physiol. **147**: 591-598.
9. **Russell, J. A., C. N. H. Long, & F. L. Engel**
1944. J. Exp. Med. **79**: 1.
10. **Engel, F. L., M. G. Winton, & C. N. H. Long**
1943. Ibid. **77**: 397.
11. **Fox, C. J., Jr.**
1944. J. A. M. A. **124**: 207.
12. **Levine, R., B. Huddleston, H. Persky, & S. Soskin**
1944. Am. J. Physiol. **141**: 209.
13. **Mylon, E., C. W. Cashman, Jr., & M. C. Winternitz**
1944. Ibid. **142**: 638.
14. **Wiggers, H. C., & R. C. Ingraham**
1946. Ibid. **146**: 431.
15. **Frame, E., & J. E. Russel**
1946. Endocrinology. **39**: 420.
16. **Engel, F. L., H. C. Harrison, & C. N. H. Long**
1944. J. Exp. Med. **79**: 9.
17. **Sayers, M. A., G. Sayers, M. G. Engel, F. L. Engel, & C. N. H. Long**
1945. Proc. Soc. Exp. Biol. & Med **60**: 20.
18. **Russell, J. A., & C. N. H. Long**
1946. Am. J. Physiol. **147**: 175.
19. **Hoar, W. S., & R. E. Haist**
1944. J. Biol. Chem. **151**: 331.
20. **Govier, W. M.**
1943. J. Pharm. & Exp. Therap. **77**: 40.
21. **Greig, M. E., & W. M. Govier**
1943. Ibid. **79**: 169.
22. **Govier, W. M., & M. E. Greig**
1943. Ibid. **79**: 210.
23. **Greig, M. E.**
1944. Ibid. **81**: 164.
24. **deTurk, W. E., & M. E. Greig**
1945. Ibid. **83**: 220.
25. **Blalock, A., & S. Levy**
1937. Am. J. Physiol. **118**: 734.
26. **Wiggers, C. J., D. F. Opdyke, J. R. Johnson**
1946. Ibid. **146**: 192.
27. **Lehninger, A.**
1945. J. Biol. Chem. **157**: 363.
28. **Sayers, M. A., G. Sayers, & C. N. H. Long**
1946. Am. J. Physiol. **147**: 155.
29. **Wilhelmi, A. E., J. A. Russell, & C. N. H. Long**
1945. Ibid. **144**: 683.
30. **Rosenthal, O., H. Shenkin, & D. L. Drabkin**
1945. Ibid. **144**: 331.
31. **Van Slyke, D. D.**
1918. Ann. N. Y. Acad. Sci. **19** (4): 593.
32. **Fine, J., H. A. Frank, & A. M. Seligman**
1945. Ann. Surg. **122**: 652.
33. **Wilhelmi, A. E., J. A. Russell, F. L. Engel, & C. N. H. Long**
1945. Am. J. Physiol. **144**: 669.
34. **Wilhelmi, A. E., M. G. Engel, & C. N. H. Long**
1946. Ibid **147**: 181.

CERTAIN ANATOMO-PATHOLOGIC ASPECTS OF HEMORRHAGE

By Paul Klemperer

Division of Pathology, Laboratories of The Mount Sinai Hospital, New York, N. Y.

The pathologist, primarily concerned with alteration of structure in disease, approaches the problem of hemorrhage with the chief question, "Where and how does blood escape from the closed channels of the circulation?" While he is aware of the significance of hemostasis for the arrest of hemorrhage, he realizes that this complex phenomenon depends primarily upon chemical and physical factors and that for its investigation physiologic rather than anatomic methods are appropriate. Yet, he will not neglect to search for morphologic evidence which might account for a disturbance in the mechanism of hemostasis.

Massive, mostly exsanguinating hemorrhages, commonly of sudden onset, are generally revealed at autopsy to be the result of perforation of the heart or of rupture of arterial or venous branches. The causative role of physical injury hardly deserves further mention, except for the reference that indirect trauma (thorax compression) may cause such disruption even without the evidence of external injury. It might also be mentioned that even large perforations of the aorta need not cause death but may heal with the formation of a false aneurism. Among the morbid states responsible for the perforation of the heart, acute myomalacia due to coronary occlusion is almost the exclusive cause. Exceptional cases of gumma or abscess of the myocardium with perforation have been reported. Rupture of the aorta or the main arteries are due either to intrinsic alterations of the wall such as syphilitic mesaortitis, or medianecrosis of the aorta, or to atherosclerosis. Perforation of smaller arteries occurs in periarteritis nodosa, mycotic aneurysms due to infected emboli or to a primary structural weakness of the wall. Perforation of varicose veins such as the dilated esophageal veins, in conditions of embarrassed portal circulation, is frequently the source of fatal hemorrhage. The erosion of arteries, less frequently of veins, due to adjacent suppuration or ulceration is responsible for massive and sudden hemorrhage. Summarizing the anatomic-pathologic experiences in these catastrophic hemorrhagic episodes, it can be stated that the site and mechanism of the blood extravasation is generally obvious and easily ascertained.

The situation is different in hemorrhagic conditions characterized by multiplicity and small dimensions of the initial hemorrhages, which by fusion may become extensive. While their size points to blood extravasation from small calibered vessels, the actual site and mechanism of the escape of blood is often not easily demonstrable. Rupture of small arterioles due to intrinsic damage of their walls or to erosion undoubtedly occurs. Structural alteration of such vessels may also produce hemodynamic disturbances in the vascular bed distal to the affected portion of the arteriole, giving rise to stasis and even hemorrhagic infarction of the area supplied. The histologic study of petechiae and ecchymoses in acute bacterial infections such as in meningococcemia, streptococcemia (subacute bacterial endocarditis), and also in Rickettsial diseases, affords the opportunity of recognizing the complicated mechanism by which arteriolar damage provokes hemorrhagic phenomena.

The escape of blood from capillaries will obviously cause minute hemorrhages. It should be kept in mind, however, that purpuric spots are not always due to actual blood extravasation but may be the result of maximal dilatation of capillaries. The differentiation can only be made by histologic examination. Hemorrhage from capillaries may be the result either of rupture or of diapedesis of red cells. Changes in the capillary wall causing fragility, or increased hydrostatic pressure, are assumed to cause rupture, while diapedesis implies the migration of erythrocytes through intact capillary walls. The histologic differentiation between these two mechanisms of blood extravasation is extremely difficult. The apparent simplicity of the structure of the capillary wall has deluded investigators into a belief that alterations could be easily ascertained. However, its integrity often cannot be determined by the conventional methods of histology. Moreover, investigation of recent years have revealed details of its intricate structure which have not yet found sufficient consideration in the microscopic study of morbid conditions associated with hemorrhagic manifestations.

It is universally accepted that the internal layer of the capillary wall is formed by continuously disposed endothelial cells. In recent years, it has been found that the endothelial cells have strong alkaline phosphatase activity. The other constituents of the capillary wall are the intercellular cement substance and an external basement membrane reinforced by a fibrillar network.[1] The latter component of the capillary wall seems to be closely related to the intercellular substance of the connective tissue, being composed of a homogeneous ground substance and fibers. The pericapillary cells do not form an integral part of the capillary wall engaged in its structural resistance and need not be considered. Some consideration, however, should be given to the elastic tissue in the immediate vicinity, because according to Peck, Rosenthal, and Erf[2] it exerts a cushioning effect and may protect the capillary wall in case of increased intracapillary pressure.

The integrity of the capillary wall may be affected by alterations in the endothelial cells, the intercellular cement substance, or the external basement membrane. In anatomic investigations of hemorrhagic conditions, these three components of the vascular wall should be submitted to microscopic scrutiny, in order to ascertain the site and mechanism of the blood extravasation. Unfortunately, adequate studies are very fragmentary, and the contribution of anatomic pathology to the problem of purpura is more conjectural than objective.

Unequivocal degenerative changes of the capillary endothelial cells are difficult to demonstrate, and the rather frequent statements in the literature about endothelial damage should be accepted with considerable reserve. Actual capillary damage initiated by endothelial injury can be demonstrated in the renal glomeruli is streptococcic infections and is considered as an adequate reason for the frequent hematuria. But the generalized hemorrhagic purpura in bacterial infections is most frequently the result of bacterial emboli in arterioles. Inadequate composition of the intercellular matrix of the supporting tissues has been established as the pathogenetic principle in human and experimental scurvy. By analogy, a similar structural defect of the basement membrane of capillaries has been postulated but not yet definitely demonstrated by histologic investigation.[3] Nevertheless, the experimental evidence of endothelial proliferation in scurvy and the subsidence of capillary fragility after vitamin C administration strongly support the hypothesis that a defect of the basement membrane and not of the endothelial cells is responsible for the hemorrhagic tendency in scurvy. Yet, a possible role of an endothelial alteration in the vascular damage still awaits further clarification, since recent investigations by Danielli, Fox, and Kodicek[4] have shown that alkaline phosphatase is strikingly decreased in the regenerating connective tissues of scorbutic guinea pigs.

Acute disseminated lupus erythematosus not infrequently presents symptoms of a hemorrhagic diathesis, evidenced by cutaneous and renal hemorrhages. Damage of the capillary basement membrane has been demonstrated as part of the systemic alteration of the connective tissues characteristic of this disease. The associated thrombocytopenia can be considered as of additional significance.

A critical review of the pathogenesis of the hemorrhagic manifestations in the conditions discussed so far, indicates that the escape of blood can be attributed to an actual alteration of the terminal branches of the circulation. An inquiry into the role of vascular fragility in the primary or secondary thrombocytopenic purpuras, however, reveals that our knowledge regarding the morphologic integrity or deficiency of the peripheral vascular bed is highly inadequate. Recent textbooks by Wintrobe[5] and Kracke[6] refer to defects in the vascular endothelium and to increased capillary permeability as fundamental factors which, in association with thrombocytopenia, are responsible for the hemorrhagic manifestation. However, there is hardly an unequivocal histologic evidence of capillary lesions which would support such a contention.

Bedson[7] believed in a toxic action of blood platelet antiserum on the endothelium of vessels evidenced by a "swollen and edematous condition of the endothelial cells." Tocantins and Stewart,[8] on the other hand, were not impressed by histologic evidence of vascular damage as a factor in the mechanism of hemorrhage in experimental thrombopenic purpura. Peck, Rosenthal, and Erf,[2] in their histologic investigations of purpura in man, found fragmentation and diminution of the elastic tissue, but no constant lesions of the capillary wall proper. It is obvious that these conflicting and fragmentary observations are inconclusive as to the significance of capillary alteration for the blood extravasation. Added histologic investigations of petechiae, using the recent techniques of phosphatase activity determination for the endothelial cells, as well as special attention to the basement membrane of the capillaries are urgently needed. In only one variety of thrombocytopenic purpura, has histologic examination revealed features which explain the mechanism of the blood extravasation. The rare instances of febrile anemia with thrombocytopenic purpura[9, 10] present generalized platelet thrombi within arterioles and precapillaries. The occlusion of the terminal arterial ramifications is associated with diapedesis of red cells from the dilated venous capillaries and venules. Clinical and histologic examination, however, have not yet been able to clarify the pathogenetic factors responsible for the precipitation of platelets within the vascular lumen. An endothelial damage has been considered but not unequivocally demonstrated. This problem, too, must await further elucidation.

It is obvious that in this short presentation of certain aspects of hemorrhage no complete survey of the problem, as it concerns the pathologic anatomist could be attempted. Many conditions, such as the allergic, endocrinal, and metabolic purpuras were completely omitted, others only indirectly referred to. The object of the essay, as indicated in the introduction, was only to discuss the site and mechanism of blood extravasation as revealed by anatomic investigation. A review of the available data indicates that massive hemorrhage has been adequately accounted for by this method, whereas the minute blood extravasation characteristic of the various clinical forms of purpura still requires further morphologic study.

BIBLIOGRAPHY

1. **Volterra, M.**
   1927. Ricerche sul sistema reticolo istiocitario lo sperimentale. Arch. di Biol. **81**: 319.
2. **Peck, S. M., N. Rosenthal, & L. Erf**
   1937. Purpura. Arch. Derm. & Syph. **35**: 831.
3. **Dalldorf, G.**
   1939. The pathology of vitamin C deficiency. *The Vitamins:* 340. A. M. A.
4. **Danielli, J. F., H. B. Fell, & E. Kodicek**
   1945. The enzymes of healing wounds. British J. Exp. Path. **26**: 367.
5. **Wintrobe, M. M.**
   1942. Clinical Pathology. Lea & Febiger. Philadelphia.
6. **Kracke, R. R.**
   1941. Diseases of the Blood. Lippincott. Philadelphia.
7. **Bedson, S. P.**
   1922. Blood platelet antiserum, its specificity and role in the experimental production of purpura. J. Path. & Bact. **25**: 94.

8. **Tocantins, L. M., & H. L. Stewart**
1939. Pathological anatomy of experimental thrombopenic purpura in the dog. Am. J. Path. **15**: 1.
9. **Moschkowitz, E.**
1925. An acute febrile pleiochromic anemia with hyaline thrombosis of the terminal arteries and capillaries. Arch. Int. Med. **36**: 89.
10. **Baehr, G., P. Klemperer, & A. Schifrin**
1936. Acute febrile anemia and thrombocytopenic purpura with platelet thrombosis in capillaries and arterioles. Trans. Assoc. Am. Phys. **51**: 43.

DISCUSSION OF THE PAPER

Dr. M. Volterra (*The Mount Sinai Hospital, New York, N. Y.*):

I wish to thank Dr. Klemperer for acknowledging my work of 1927 on the structure of the capillary wall. Since that time, I have emphasized the importance of the reticular adventitia, or basal membrane, in relation to the resistance of blood capillaries. I do not think that damage to the endothelial cells is a primary cause of capillary fragility. It is a pity that more attention has not been paid, by the pathologist, to possible alteration of the basal membrane of the small vessels in cases of purpura. The failure is probably due, as Dr. Klemperer has pointed out, to the fact that the technique for the study of those structures is very difficult. I would like to add that the majority of the silver impregnation methods used to that end, do not show the thin layer of amorphous substance which, normally, is present between the meshes of the fibrillar net and forms a continuous surface externally to the endothelium. Drs. Rosenthal and Peck have, in the past, referred to possible alteration in the elasticity of the small vessels. Blood capillaries do not have elastic connective tissue fibers around the outside walls, the elasticity being a property of the reticular fibers of the capillary basal membrane.

HEMORRHAGIC MANIFESTATIONS OF BACTERIAL AND VIRUS INFECTIONS: EXPERIMENTAL STUDIES AND PATHOLOGICAL INTERPRETATIONS

By GREGORY SHWARTZMAN AND I. E. GERBER
Division of Bacteriology, Laboratories of the Mount Sinai Hospital, New York, N. Y.

During recent years, there has accumulated a considerable amount of experimental work on various aspects of the phenomenon of local tissue reactivity. The limited time allowed for this presentation does not permit adequate mentioning of the important contributions of many investigators. A complete review of the literature up to 1937 will be found in a monograph on the subject.[1]

Here, an attempt will be made to point out some correlations of the studies on the phenomenon, with the localized and disseminated hemorrhagic manifestations superimposing or occurring incidentally to certain spontaneous and experimental bacterial and virus infections.

Although the basic experimental observations are well known, it will be necessary to begin the presentation by a restatement of the initial experiment, as presented in TABLE 1. It will certainly not be our objective to offer a comprehensive description of the various aspects of the phenomenon. We shall limit ourselves to an outline of the salient characteristics of the phenomenon which have to be kept in mind in connection with the interpretations to follow later.

TABLE 1

<table>
<tr><th colspan="3">Basic Experiment</th></tr>
<tr><th>Local Preparation</th><th></th><th>Elicitation</th></tr>
<tr><td>Intradermal preparatory injection of active bacterial filtrate</td><td>24 hours later</td><td>Intravenous provocative injection of active bacterial filtrate</td></tr>
<tr><th colspan="3">Reaction</th></tr>
<tr><td colspan="3">Four hours following intravenous injection: severe hemorrhagic necrosis at the site of the intradermal injection.</td></tr>
</table>

The first salient feature to be taken up are the gross appearance and histological changes in the site of the reaction.

Grossly, the striking feature of the phenomenon is a rapid and violent hemorrhagic necrosis with sharply demarcated borders at the prepared site following the intravenous provocative injection. Although the characteristic lesion is clearly demarcated grossly, there exists evidence that it is more extensive than it may appear at first inspection. The toxin injected intradermally is picked up by the lymphatics and thus conveyed to the regional lymph nodes. Following the intravenous provocative injection, a reaction is obtained in these lymph nodes which is of the same character and severity as in the prepared skin site itself.[2]

The microscopic studies on the reaction in the skin and lymph-nodes (PLATE 3) confirm the gross observation that hemorrhagic necrosis is an event of primary importance in the phenomenon under consideration. Studies on the histology of the reaction in the skin sites at various time-intervals following the intravenous provocative injection, indicate that the tissue changes present themselves in the following order:

There first ensues a marked degree of venous and capillary dilatation and engorgement. This is immediately followed by severe hemorrhage. Shortly thereafter, there occurs a striking degree of edema and intense leucocytic infiltration. At this time, many veins and capillaries are filled with thrombi consisting of platelets with enmeshed leucocytes, erythrocytes, and fibrin strands. The thrombi completely fill the lumens, while in larger vessels they are parietal. This picture is accompanied by large central areas of necrosis.

While the tissues surrounding the site of the reaction are histologically normal, yet certain macroscopically not detectable changes take place which express themselves in enhanced fragility of the capillaries, a manifestation resembling closely the so-called capillary fragility observed in certain conditions of hemorrhagic diathesis. This is demonstrated as follows:

Following the intravenous provocative injection when the local reaction is clearly demarcated and fully developed, the application of stimuli, such as suction, application of a clamp, pulling of hair, shaving, and the like brings about the appearance of petechiae in histologically normal adjacent tissues.

When the phenomenon is elicited in rabbit's ear by the combined intradermal injection into one ear and intravenous injection into another, portions of the ear not affected by the reaction frequently become cyanotic. It is worthy of special note that, in these sites, it is possible to detect, microscopically, thrombosis of the venules without any perivascular reaction.

No significant blood alterations could be held responsible for the state of experimental hemorrhagic diathesis as shown by complete hematological studies made on sixty-one normal and experimental male and female rabbits. No changes associated with the phenomenon could be observed in the bleeding, coagulation, clot retraction, and sedimentation times. Leucopenia and a drop in hemoglobin and erythrocyte count, encountered in some animals, had no relation to any phase of the phenomenon. Although the intravenous injection of meningococcus filtrate gave a marked decrease in the number of platelets, positive animals receiving other bacterial filtrates nevertheless showed normal counts.

Thus, it is possible to elicit, by means of certain bacterial filtrates and under certain conditions, a state of capillary fragility. The state is not expressed by morphological endothelial alterations, or by any blood changes; and yet, there occurs some profound functional disturbance which leads to a tendency to bleeding and capillary thrombosis. The state of experimental hemorrhagic diathesis described in connection with the phenomenon closely resembles certain hemorrhagic manifestations encountered at times during the course of bacterial and virus infections in man. Here, no consistent evidence is found of capillary damage or disturbance in the clotting mechanism, namely, the symptomatic hemorrhagic diathesis. This relationship will be discussed later in greater detail.

The second feature of the phenomenon concerns the routes for preparatory and provocative injections. In the basic experiment, the preparatory injection is made into the tissue, while the provocative injection is only effective by way of the general circulation. In a later modification of the basic experiment, it was shown that the preparatory injection could also be made by way of the general circulation. When one intravenous injection of an active bacterial filtrate was followed, twenty-four hours later, by a second intravenous injection of an active filtrate, typical lesions of the phenomenon were obtained in the internal organs, the most striking ones being revealed in the kidneys. This manifestation was termed "general phenomenon" of tissue reactivity.

The view is generally accepted, based on unequivocal evidence, that the phenomenon under discussion is not anaphylactic in nature, namely, that the state of preparation resulting from the intradermal injection of the filtrate does not represent an anaphylactic sensitization. It is certainly impossible to discuss the evidence within the space allowed for this paper. It may be mentioned, however, that this realization, early in the work, led to the detection of the non-specificity of the phenomenon, a fact which, as it will be pointed out later, represents a feature useful in the interpretation of certain clinical processes.

The non-specificity of the phenomenon is illustrated in TABLE 2. With this observation as a basis, obviously a great variety of bacterial and non-bacterial factors were examined for their ability to produce the phenomenon. In this manner, some relationships (presented in TABLES

TABLE 2

NON-SPECIFICITY OF PHENOMENON OF LOCAL TISSUE REACTIVITY

Preparation: Intradermal injection	Elicitation: Intravenous injection
Meningococcus filtrate (A)	*Meningococcus* filtrate (A)
B. typhosus filtrate (B)	*B. typhosus* filtrate (B)
B. coli filtrate (C)	*B. coli* filtrate (C)
A	A, B, or C
B	A, B, or C
C	A, B, or C

3 and 4) came to light. Although appearing rather intricate, the relationships are clearly demonstrated if the essential criteria for recognition of the phenomenon are kept in mind, namely: the type of lesion, speed of its appearance, the necessity to give the provocative injection by way of the general circulation and the short incubation period necessary for obtaining the state of preparation.

TABLE 3

FACTORS CAPABLE OF ELICITING THE PHENOMENON

Local preparatory injection	Intravenous provocative injection	Reaction
Non-bacterial, irritating inflammatory and toxic substances	Active bacterial filtrate	Negative
Active bacterial filtrate	Active bacterial filtrate	Positive
Vaccinia virus infection	Active bacterial filtrate	Positive
Local bacterial infection: *B. anthracis, B. tuberculosis, Streptoc. hemolyticus, Pneumococcus friedländeri, H. influenzae,* etc.	Active bacterial filtrate	Positive

In TABLE 3, there are indicated the agents which are capable of inducing the necessary state of reactivity. Obviously, the ability is strictly limited to bacteria and their products and vaccinia virus. Conversely, as may been seen from TABLE 4, the group of agents capable of provoking the reaction in the prepared sites includes, in addition to bacterial products, certain other substances and antigen-antibody complexes. The typical experiment for elicitation of the reaction to the complex of an animal protein antigen with actively acquired antibodies is given in TABLE 5. When considered together, the features just described (summarized in TABLE 6) suggest certain pathological and clinical interpretations, to be discussed here at length.

As previously noted, the phenomenon of tissue reactivity may be produced in the internal organs of rabbits under certain conditions, namely

TABLE 4

FACTORS CAPABLE OF ELICITING THE PHENOMENON

Local preparatory injection	Intravenous provocative injection	Reaction
Animal and plant proteins, agar, starch, and glycogen	Active bacterial filtrate	Negative
Active bacterial filtrate	Agar, starch*	Irregular
Active bacterial filtrate	Horse serum	Negative
Active bacterial filtrate	Horse serum—rabbit anti-horse serum (*in vitro* mixture)	Positive
Active bacterial filtrate	Horse serum injected intravenously into rabbit previously sensitized by horse serum	Positive

* Bier & do Amaral[8] recently reported that they could successfully provoke the phenomenon in skin sites prepared with a meningococcus culture filtrate by means of the intravenous injection of glycogen. The substance was, however, ineffective as a preparatory agent.

TABLE 5

EFFECT OF ANTIGEN-ANTIBODY COMPLEXES UPON TISSUES PREPARED WITH BACTERIAL PRODUCTS

(1) Rabbit sensitized with horse serum by a single injection by any route.
(2) One week later, active bacterial filtrate is injected intradermally.
(3) Twenty-four hours later, horse serum is injected intravenously.

Reaction:

Four hours later, there appears typical hemorrhagic necrosis at the site prepared by the intradermal injection of bacterial filtrate.

TABLE 6

SUMMARY: FACTORS CAPABLE OF ELICITING THE POSITION OF LOCAL TISSUE REACTIVITY

Local state of preparation	Elicitation
Certain bacterial products and bacterial and virus infections	Active homologous and heterologous bacterial products; antigen-antibody complexes and live bacteria

by two intravenous injections of bacterial filtrates at an interval of twenty-four hours. Similarly, viable bacteria may be used as a preparatory factor. In such instances, a single intravenous injection of 3 cc. of an eighteen-hour-old 0.3% glucose broth culture of *Streptococcus viridans* is followed, forty-eight or seventy-two hours later, by an intravenous injection of meningococcus or other potent bacterial filtrate. The visceral lesions produced are similar, irrespective of whether viable bacteria or bacterial filtrates are used as preparatory factors. These consist, in the main, of renal lesions of variable character and extent, such as focal glomerular capillary thrombi, often with associated glomerular hemorrhages, tubular necrosis, and focal or extensive cortical necrosis with marked vascular involvement and massive hemorrhage. Simultaneously, mural vascular thrombi are seen in the liver, spleen, bone marrow, and lungs, and occasionally adrenal necrosis is observed.[4]

Quite comparable lesions have been noted in the kidneys of horses repeatedly injected with meningococcus filtrates for the production of immune sera, and in spontaneous hog cholera and swine erysipelas. Similar vascular alterations are encountered in man, often in association with hemorrhagic manifestations, the so-called purpuras of infectious origin. Such hemorrhagic manifestations have usually been ascribed to capillary injury, said to arise during the course of a bacterial or virus infection, although significant histological evidence of such injury is lacking.

It seems possible to interpret the abovementioned pathological alterations on the basis of the phenomenon of tissue reactivity. For example, in meningococcemia, the bacteria are often demonstrable on direct smear or culture of the purpuric lesion. It is likely that the bacteria at the local site prepare the tissues and that the hemorrhagic reaction is then elicited by the meningococcus toxins which circulate in the blood stream. Such hemorrhages may at times be massive and are often accompanied by tissue necrosis. The well-known hemorrhagic and destructive adrenal lesion seen in meningococcus infection, producing the Waterhouse-Friderichsen syndrome, may also be accounted for by this mechanism. The above conclusions are supported by the direct evidence that the phenomenon under discussion may be encountered in man.

There are many cases on record of sudden death following intravenous injections of typhoid vaccine. The most recent ones, namely, those reported by Urbach, Goldburgh, and Gottlieb[5] and by Love and Driscoll[6] were carefully examined by the authors, in order to determine their relation to the phenomenon of local tissue reactivity. Thus, in the first case the following events took place:

For non-specific therapy, the patient received an intramuscular and 30 minutes later an intravenous injection of 10 and 50 million typhoid bacilli, respectively. No reaction occurred. Twenty-four and forty-eight hours later, the same injections were repeated. The last injection was followed, six hours later, by a severe reaction and death. *Post-mortem* examination revealed numerous purpuric spots of the skin, pericardium, endocardium, and liver surface. Kidneys were studded with hemorrhagic areas, and the lungs were acutely congested, showing localized hemorrhagic lesions on subpleural surfaces. Histologically, there was tubular necrosis of the kidneys, acute pulmonary congestion, hemorrhage in the adrenals, and acute hepatic necrosis. There was nothing in the history of the patient to suggest any purpuric tendency. The second case (Love and Driscoll's) was similar. The conditions under which the reactions were obtained were clearly those of the general phenomenon already mentioned, in which both preparatory and provocative injections are given intravenously. These cases are accidental but clear-cut demonstrations of the phenomenon in man, indicating that the experimental basis set forth in laboratory animals is justifiably applicable to clinical interpretations.

There is also certain indirect evidence that toxins which are capable of producing the phenomenon of tissue reactivity in animals may be responsible for the pathologic alterations just described. For example, meningococcus toxins capable of producing the phenomenon in rabbits may be neutralized by the antibodies which develop in man during the course of the disease, as shown by Powell and Jamieson[7] and by Klein.[8] Stolyhwo[9] also showed that principles capable of producing the phenomenon in animals may be isolated from the urine of highly toxic cases of typhoid fever.

It has already been mentioned that the state of local and general reactivity is observed not only following the injection of bacterial products, but also in actively infected tissues. In the experiments dealing with the local tissue preparation, live bacterial suspensions were injected intradermally. After various intervals of time when the local lesion showed no hemorrhagic necrosis, bacterial products or live microorganisms were injected intravenously. Severe hemorrhagic lesions appeared in the infected sites.

In the studies of Gratia and Linz,[10] Bordet,[11] Witebsky and Neter,[12] Cohen and Moolten,[13] and in my own investigations,[1] the infecting organisms were *Bacillus anthracis, Bacillus friedländeri, Pneumococcus, Streptococcus hemolyticus* and *viridans, Hemophilus influenzae,* gram-negative anaerobes isolated from pulmonary abscess of man, and *Bacillus tuberculosis,* strain BCG. The latter organism was especially suitable for these studies, since it produces a low-grade, slowly developing lesion The significant fact in this connection is that the state of preparation which exists only for some twenty-four hours when filtrates are used, may persist for several weeks in BCG-infected sites.

Obviously, these observations serve to demonstrate the effect exercised by a bacterial invasion of the blood, or by toxemia, upon the course of local infections. They also suggest the pathogenesis of sudden transformations of mildly inflammatory reactions into hemorrhagic and necrotic lesions seen in such infections as appendicitis, pancreatitis, and possibly hemorrhagic gastritis.[14] The hemorrhagic manifestations in tuberculous lesions following intercurrent infections also seem to belong to this category. The hemorrhagic forms of smallpox are of special interest here. As is well known, there are two clinical hemorrhagic forms of the disease. Both are almost always associated with a bacterial septicemia, especially *Streptococcus hemolyticus.* The hemorrhages may diffuse into the corium preceding the visible eruption, or, at a later stage, localize in the pustules. Both forms of hemorrhage may be readily produced experimentally. Thus, in the experiments of Gratia and Linz,[10] vaccinia virus was rubbed into the skin of a rabbit. Forty-eight hours later, before any eruption occurred, the intravenous injection of an active bacterial filtrate transformed a slight erythema into a severely hemorrhagic and necrotic lesion. Curiously, the lesion changed in this manner went on to healing without developing into a pustule. In another group of experiments the virus was injected intratesticularly, followed, forty-eight hours later, by the intravenous injection of the active bacterial filtrate. Four to twenty-four hours later, the transformation of the vaccinia lesion of the testicle was extremely pronounced. The testicle showed a violent hemorrhagic reaction, extending along the epididymis and spermatic cord. The peritoneal cavity was frequently filled with blood. Lymph nodes, bladder, and intestines were severely hemorrhagic.

PLATES 4 and 5 illustrate the histological changes taking place in the testicles, as confirmed by Koplik in this laboratory.

These observations on combined effect of a virus and bacterial invader or its products are consistent with the growing belief that some specific clinical entities may be caused by the concerted effect of two biologically unrelated etiological agents. In human spontaneous influenza, in Shope's experimental studies on swine influenza, and in human influenza virus in mice,[15] the secondary bacterial infection contributes to the variable pathology of the lungs. The modifications of virus-produced lesions by means of the bacterial agents are complex and not sufficiently clear. It seems, however, that at least in some forms, the pulmonary hemorrhagic manifestations are conspicuous. In 1937, Watjen examined 57 fatal cases.[16] Confirming well-known previous studies of McCallum, Goodpasture, and many others, he found that widespread hemorrhagic involvement was a conspicuous feature in lungs yielding *Streptococcus hemolyticus*. Thus, there is suggestive evidence that the pathogenesis described in connection with hemorrhagic vaccinia infection may also be operative in influenza. Experimental studies in this direction are in progress.

As previously emphasized, tissues prepared by live organisms or bacterial products respond with a severe hemorrhage not only to intravenous injection of bacterial products but to the complexes of unrelated proteins and specific antibodies. Histologically, this reaction is identical with the one already described in the basic experiment. The experiments demonstrate a relationship between bacterial infections and anaphylactic sensitization to non-bacterial proteins. Accordingly, it may be expected that an anaphylactic reaction may be significantly modified by an intercurrent infection. Thus, the local anaphylaxis, the Arthus phenomenon in man, usually runs a mild course. As seen from a report on a very large number of cases by Lucas and Gay,[17] a protein injection into the skin of a patient sensitized to the protein, results in most cases in edema and infiltration. The reaction may, however, appear extremely severe when elicited in a patient suffering from some infectious disease. In the cases of Irish and Reynolds[18] and of Kohn, McCabe, and Brem,[19] the reactions were violent, resulting in extensive hemorrhage and necrosis over a large area of the buttocks. There were clear-cut indications of intercurrent distant infections in these cases, otitis media and membraneous pharyngitis, respectively. The local lesion was sterile on repeated cultures.

To the same category may belong the so-called "hetero-allergic flare-up" of dormant infections elicited by the injection of non-bacterial antigens into allergic individuals. Experimentally, Dienes[20] produced severe hemorrhagic lesions in tuberculous foci by the injection of egg albumen into guinea pigs previously sensitized by this antigen. Local hemorrhagic-necrotic skin reactions infrequently seen in allergic individuals in association with bacterial infections also deserve consideration from this point of view (Harkavy and Romanoff[21]).

Thus, it appears that the phenomena of local and general reactivity permit the interpretation of certain pathological sequellae of human bacterial and virus infections as the result of tissue reactivity to the infecting agents or their toxins produced *in vivo* in the course of active infection. These pathological alterations are manifested as local or general hemorrhages, vascular thrombi and necrosis, and occasionally as tissue necrosis. The factors which predispose to the occurrence of these changes in any given case are not completely known and are undoubtedly quite variable, since there are wide variations in individual response, in the phenomenon-producing properties of bacterial strains, and in the conditions under which elicitation of the phenomenon may be obtained.

The interpretations given may be variously evaluated, being considered conclusive in some instances and only suggestive in others. A certain general conclusion seems, however, to be of greater significance. The concept of mono-etiology dominates our approach to investigations of infectious syndromes. It seems that sufficient evidence has been accumulated to point out the importance of investigating the possibility that some diseases may be caused by the combined effects of plural etiological agents, biologically remote from one another, as viruses and bacteria, or bacteria and complexes of non-bacterial protein antigens with specific antibodies. The synergism of these agents, under suitable conditions, may thus result in characteristic manifestations altogether different from the effects of each agent acting independently.

BIBLIOGRAPHY

1. **Shwartzman, G.**
   1937. Phenomenon of Local Tissue Reactivity. Paul B. Hoeber (Harper & Bros.). New York.
2. **Koplik, L. H.**
   1937. Experimental production of hemorrhage and vascular lesions in lymph nodes: an extension of the Shwartzman phenomenon. J. Exp. Med. **65**: 287.
3. **Bier, O. G., & J. P. do Amaral**
   1944. Desencadeamento do fenomeno de Shwartzman em coelhos mediante a injeção venosa de glicogenio puro. Mem. Inst. Butantan **18**: 33.
4. **Gerber, I. S.**
   1936. The Shwartzman phenomenon in the kidneys of rabbits. Arch. Path. **21**:776.

   Shwartzman, G., A. J. Bernheim, & I. S. Gerber
   1938. The phenomenon of tissue reactivity in the kidneys of rabbits. J. Mount Sinai Hosp. 4: 1020.
5. **Urbach, E., H. L. Goldburgh, & P. M. Gottlieb**
   1944. General Sanarelli-Shwartzman phenomenon with fatal outcome following typhoid vaccine therapy. Ann. Int. Med. **20**: 989.
6. **Love, J., & R. H. Driscoll**
   1945. Anaphylactoid (Sanarelli-Shwartzman) reaction following therapeutic antityphoid injections, with report of fatal case. U. S. Naval Med. Bull. **45**: 1104.
7. **Powell, H. M., & W. A. Jamieson**
   1931. Antimeningococcic serum with special reference to the Shwartzman phenomenon. Am. J. Hyg. **14**: 470.
8. **Klein, H. M.**
   1933. Studies on the Shwartzman phenomenon: hitherto undemonstrated antitoxins in man. J. Inf. Dis. **53**: 312.

9. **Stolyhwo, N.**
   1931. Toxische Erscheinungen im Verlauf des Typhus abdominalis und des Shwartzmanschen Phänomen. Bull. Soc. Biol. Lettonie **2:** 129.
   1936. Toxic substances in urine and sweat of typhoid fever patients as demonstrated by the Shwartzman phenomenon. J. Immunol. **30:** 235.
10. **Gratia, A., & R. Linz**
   1932. Les phénomènes de Sanarelli et de Shwartzman ou l'allergie hemorrhagique. Ann. Inst. Pasteur **49:** 131.
11. **Bordet, P.**
   1936. Contribution a l'étude de l'allergie. Ann. Inst. Pasteur **51:** 357.
12. **Witebsky, E., & E. Neter**
   1938. Pathogenesis of hemorrhagic-necrotic skin lesions in intradermal infection of rabbits with pneumococci. Proc. Soc. Exp. Biol. & Med. **38:** 187.
13. **Cohen, J., & S. E. Moolten**
   1943. The Shwartzman phenomenon in the genesis of pulmonary abscess. Arch. Path. **35:** 517.
14. **Crohn, B., & G. Shwartzman**
   1937. Ulcer recurrences attributed to upper respiratory tract infection: a possible illustration of the Shwartzman phenomenon. Am. J. Dig. Dis. Nutrit. **11:** 705.
15. **Shope, R. E.**
   1937-38. Recent knowledge concerning influenza. Am. Int. Med. **11:** 1.
   Straub, M.
   1937. Microscopical changes in the lungs of mice infected with influenza virus. J. Path. Bact. **45:** 75.
   1940. Histology of catarrhal influenza bronchitis and collapse of the lung in mice infected with influenza virus. Ibid. **50:** 31.
   Mulder, J.
   1940. *H. influenzae* and influenza virus in relation to bronchitis. Ibid. **50:** 317.
16. **Watjen, J.**
   1937. Pathologische anatomische Erfahrungen bei der Grippenepidemie des letzten Winters mit besonderer Berücksichtigung der Influenzabazillenbefunde. Dt. Med. Wochenschr. **63:** 993.
17. **Lucas, W. P., & F. P. Gay**
   1909. Localized anaphylactic intoxication in children following the repeated injection of antitoxin. J. Med. Res. **15:** 251.
18. **Irish, H. E., & E. C. Reynolds**
   1933. Massive anaphylactic gangrene. J. A. M. A. **100:** 490
19. **Kohn, J. L., E. J. McCabe, & J. Brem**
   1938. Anaphylactic gangrene following administration of horse serum (Arthus or Shwartzman phenomenon). Am. J. Dis. Childr. **55:** 1018.
20. **Dienes, L.**
   1929-30. Hemorrhagic reactions in tuberculous lesions and skin tests during protracted anaphylactic shock. Proc. Soc. Exp. Biol. & Med. **27:** 690.
21. **Harkavy, J., & J. Romanoff**
   1939. Local hemorrhagic-necrotic skin reactions in man (Shwartzman phenomenon). J. Allergy **10:** 566.

DISCUSSION OF THE PAPER

Dr. Glenn H. Algire (*National Cancer Institute, Bethesda, Maryland*):

Will you discuss the chemical nature of the active agent in bacterial filtrates, in view of the divergent views as to its protein or polysaccharide composition?

Also, do you have an explanation for the fact that a preparatory dose of bacterial filtrate is not needed to elicit hemorrhage and necrosis in certain transplanted tumors?

As to the vascular mechanism of the reaction to injection of bacterial filtrates, I should like to mention some *in vivo* observations made on transplanted sarcomas growing in transparent chambers inserted into mice. While some vessels show engorgement, as mentioned by Dr. Shwartzman, the overall picture, based on observation and quantitative measurements of the vascular bed, is one of vascular occlusion. This decreased capillary blood supply led, in tumors, to hemorrhage at the end of four hours after intraperitoneal injection of the bacterial filtrate. Capillaries in a normal tissue (striated muscle) showed a similar decrease in functional vessels and occlusion of many channels, but hemorrhage did not occur.

Dr. Gregory Shwartzman:

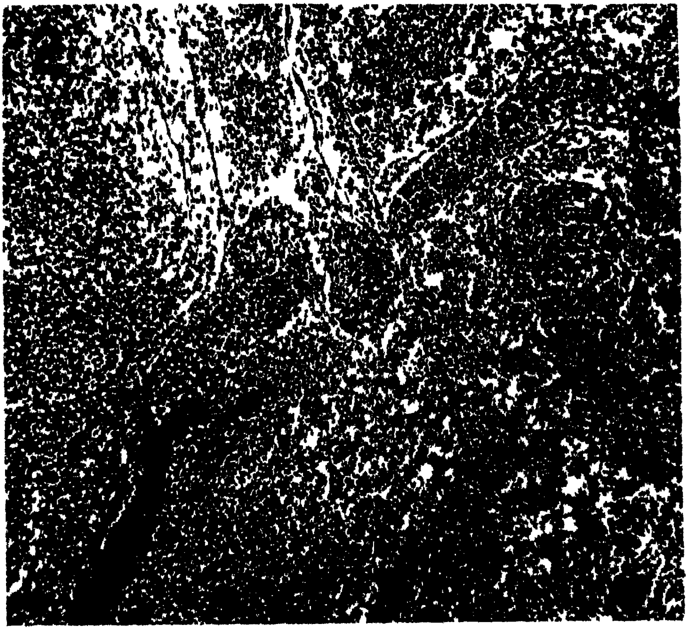
Shear and co-workers identified a polysaccharide in a concentrate of *Serratia marcescens (Bacillus prodigiosus)* filtrate. Since the active principle resisted tryptic digestion, the authors ascribed the potency to the polysaccharide. Nevertheless, the evidence supplied by other investigators, including my own, favors the assumption that the active principle is protein in nature, possibly associated with a polysaccharide. There are many protein antigens which are resistant to tryptic digestion. As a matter of fact, the trypsin failed to induce any measurable alteration in the antigenicity of the material prepared by Shear, as demonstrated by the ability of the material to stimulate production of precipitating and neutralizing antibodies in rabbits. The antibodies reacted with the concentrate as well as with the parent filtrates which served for preparation of the concentrates.

It was shown by many investigators that the phenomenon-producing principles are closely related to, or identical with, the factors capable of inducing hemorrhage and necrosis in tumors. The state of reactivity of the phenomenon in the rabbit, which results from a preparatory injection, apparently exists spontaneously in the tumor tissue. No certain explanation exists for this fact. The assumption which has been voiced but not proved is that the etiological agent of the tumor is capable of inducing the necessary state of reactivity.

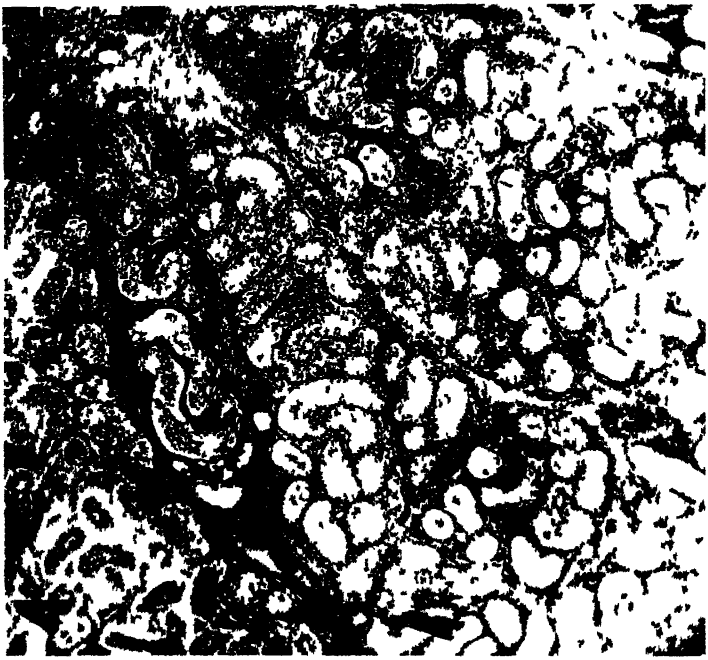
Dr. Algire's observations, made by means of the transparent-chamber method developed by him, are very interesting. It is important to extend the studies to the phenomenon in the rabbit. Histological observations on the reaction indicate that hemorrhage is a sudden event which begins to take place in a few minutes following the injection of the provocative factors. The vascular occlusion is not the direct cause of hemorrhage. Dr. Algire's observations indeed support this contention, inasmuch as the vascular occlusion occurring in normal tissue, as he just mentioned, fails to induce the hemorrhage. A state of capillary vulnerability in the reactive tissues appears to be of primary significance.

PLATE 3

Rabbit was prepared by injection into the skin of the abdomen of 0 25 cc. of diluted 1:2 filtrate of "agar washings" of meningococcus, strain 44 B. Twenty-four hours later, the rabbit received intravenously 1 cc. per kgm. of body weight of the same filtrate diluted 1:25. FIGURE 1 (x 22) and FIGURE 2 (x 100) show sections of lymph node in the vicinity of a severe local skin reaction, four hours following intravenous injection of the filtrate Note severe hemorrhagic necrosis in FIGURE 1. FIGURE 2 demonstrates engorged venules with mural thrombi and extensive hemorrhage into lymphatic sinusoids.



SHWARTZMAN AND GERBER: HEMORRHAGE AND INFECTIONS



SHWARTZMAN AND GERBER: HEMORRHAGE AND INFECTIONS

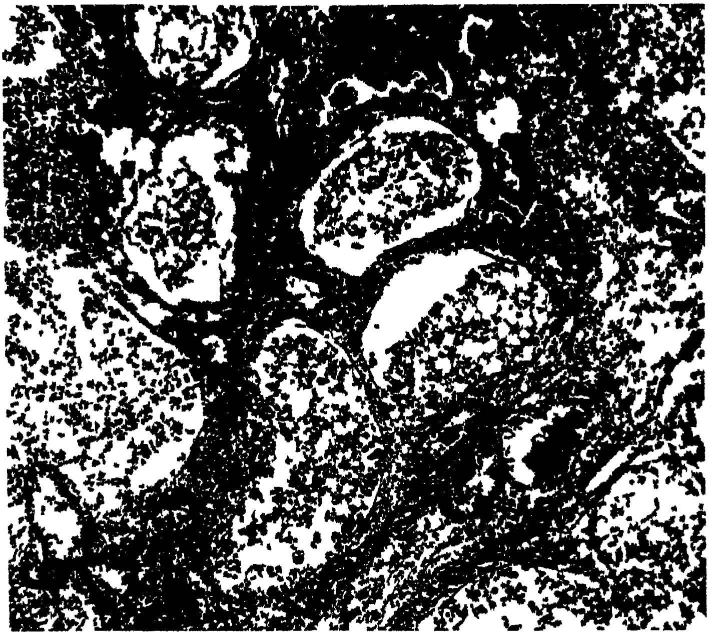
Plate 4

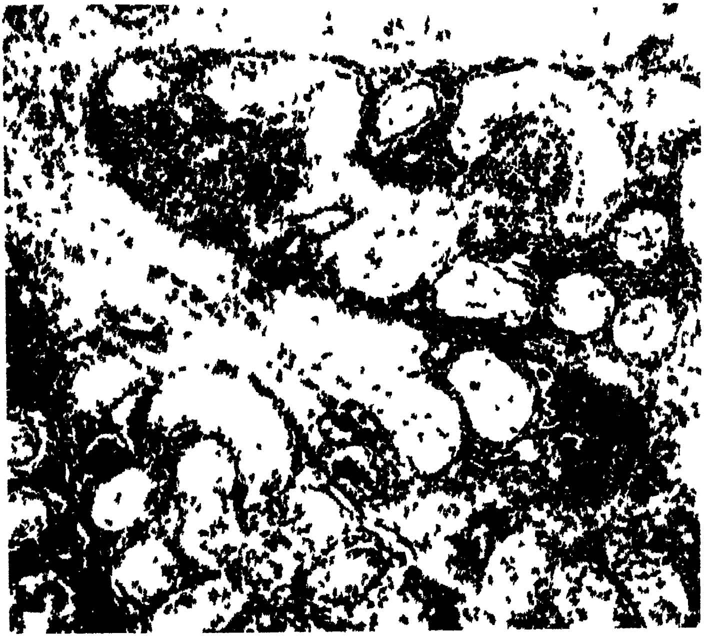
Figure 3 (× 100) Section of rabbit testicle 48 hours following intratesticular injection of testicular vaccinia virus emulsion diluted 1 100 Note capsular inflammation and edema of tubules

Figure 4 (× 22) Sections of rabbit testicle The testicle was injected with testicular vaccinia virus emulsion, diluted 1 400 48 hours later, the rabbit received an intravenous injection of 1 cc per kgm of body weight of meningococcus, strain 44 B, "agar washings" filtrate, diluted 1 25 The testicle was removed four hours following the intravenous injection Figure 4 shows extensive hemorrhagic necrosis

Plate 5

Figures 5 (x 50) and 6 (x 100) Same as on Plate 5 Note pronounced intratubular hemorrhage and necrosis of the tubules





SHWARTZMAN AND GERBER: HEMORRHAGE AND INFECTIONS

ABNORMAL HEMORRHAGE WITH NORMAL PLATELET COUNT AND NORMAL CLOTTING

By Russell L. Haden, R. H. Schneider, and Lee C. Underwood
Cleveland Clinic, Cleveland, Ohio

A patient who bleeds without apparent cause, or excessively from a minor injury, has hemorrhagic disease, for which there are numerous possible causes. The clinician must make a correct differential diagnosis if proper treatment is to be carried out.

Hemorrhage may occur in the skin and subcutaneous tissues, from mucous membranes, from point of trauma and into body cavities or joints. Red cells should not pass the endothelial barrier of an intact vessel. If a vessel wall is injured, a clot normally should be quickly formed and an abnormal loss of blood prevented.

Abnormal hemorrhage can occur only when (1) the permeability of the vessel wall is increased or (2) the clotting mechanism is disturbed. Increased permeability of the vascular endothelium allows blood to escape from the capillaries. This is the most important factor in clinical hemorrhagic disease.

Table 1

Classification of Hemorrhagic Disease (527 Patients)

Hemorrhagic Disease with *Decreased* Platelets or Abnormal Clotting Factors	
(A) Purpura due to thrombocytopenia (primary and secondary)	98
(B) Hemophilia	22
Total	120
Hemorrhagic Disease with *Normal* Platelet Count and Clotting Factors	
(A) Purpura without thrombocytopenia	105
(B) Congenital thrombasthenia	3
(C) Abnormal gross bleeding without cause	247
(D) Miscellaneous	52
Total	407

Five hundred and twenty-seven patients (table 1) with abnormal hemorrhage, seen at Cleveland Clinic during the past few years, were analyzed. Each patient had a thorough laboratory study from the standpoint of hemorrhagic disease as well as a complete history and clinical examination. Less than one-fourth (120 patients) had any significant alteration in platelets or clotting factors. Over three-fourths bled abnormally without reduction of platelets below the critical level, and without significant change in coagulation. In every patient, the abnormal hemorrhage was an important part of the clinical problem and usually the presenting complaint. Many patients had gross bleeding. Menorrhagia for which no local or glandular cause could be found was a frequent reason for making a special examination for hemorrhagic disease. Gross bleeding from the genito-urinary tract without a demonstrable local lesion, abnormal blood loss from a wound after an operative procedure, nasal hemorrhage, generalized subcutaneous ecchymoses, retinal hemorrhages are all examples of excessive bleeding of which a special study was made. In other cases, purpura was the presenting symptom.

We have a uniform method of examination when a patient presents himself at the laboratory with abnormal hemorrhage. This examination consists of:

(1) A routine blood count, hematocrit reading, and examination of a stained film.

(2) Coagulation time by the Lee-White method. Normal blood coagulates by this method in five to eight minutes. Coagulation time of over ten minutes is abnormal.

(3) Platelet count, using the Rees-Ecker technique. A normal count should be 250,000 to 300,000. A count below 200,000 is definitely reduced, although not sufficiently so to produce bleeding.

(4) Bleeding time by the methods of Duke or Ivy.

(5) Test for clot retraction.

(6) Special tests, as determination of fibrinogen or the clotting time of recalcified plasma, if indicated.

In the clinical examination, a positive tourniquet (Leede-Rumpel) test is often present, though not to such a marked degree as is constantly found in thrombocytopenic purpura. In making this test, Quick[1] suggests that a blood pressure cuff be placed on the upper arm and the pressure maintained midway between systolic and diastolic pressure for eight minutes. A circle 5 cm. in diameter is drawn on the flexor surface of the forearm with the center 4 cm. below the bend of the elbow. After fifteen minutes the petechial spots in the circle which are visible to the naked eye are counted. The upper limit of normal is +10 petechial spots.

In this large group of patients who bled abnormally without significantly reduced platelets or abnormal clotting, the primary fault is above all an increased permeability of the endothelium lining the capillaries. The vascular permeability is increased because the endothelial lining is injured. The damage may be to endothelial cells, to the cement which binds the endothelial cells together, or a combination of the two factors. This injury may occur in a number of different ways:

Infection. Many infections are associated with abnormal hemorrhage seemingly due to damage to endothelial cells by the soluble toxins produced by the infecting organism. At the onset of *meningococcus* meningitis or typhus fever, for example, purpura nearly always occurs. Almost any serious infection may, at times, cause abnormal bleeding from vascular injury.

Impaired nutrition. In scurvy, there is a deficiency of vitamin C (ascorbic acid), which is necessary for the formation of the intercellular cement substance binding the endothelial cells together. Other nutritional factors may also be concerned.

Substances having toxic effect on endothelium. Many drugs have this action. Quinine is a good example, but many other drugs act similarly, especially when the patient is sensitive to the drug in question. Snake venoms have an almost selective action on the capillary endothelium. Metabolic poisons, such as develop in kidney failure, have a similar action.

Allergic states. In allergy, one of the characteristic reactions is transudation through vessel walls. The hive, the wheal, and angioneurotic edema are all due to the escape of fluid from the blood plasma into surrounding tissue. Red cells may similarly go through vessel walls.

Anoxemia. A decrease in oxygen supply affects vascular permeability, probably through a toxic effect on endothelium.

In all such conditions, there is either a defect in the intercellular cement-substance or chemical damage to endothelial cells. Zweifach[2] thinks the most important factor in increased permeability to be the intercellular cement. This he was able to alter by changing the pH and calcium content of the fluid bathing the endothelial cells. He concludes that the physical state of the capillary wall depends on the calcium content and acidity of the perfusate, and on the endothelial colloid coating. The endothelial cells secrete the intercellular cement. Their integrity, therefore, is likewise important. The filtration barrier can thus be divided into two constituents, the intercellular cement and an absorbed layer of protein on the endothelium which can be affected independently or concurrently by the fluids on either side of the capillary.

TABLE 2

BLOOD FINDINGS IN PATIENTS WITH ABNORMAL BLEEDING NOT DUE TO THROMBOCYTOPENIA OR HEMOPHILIA

Type of bleeding	Number (total-407)	Prolonged coagulation time (10+ min.)	Prolonged bleeding time (4+ min.)	Decreased platelets	Clot retraction
Purpura					
Rheumatic (Henoch-Schönlein's)	25	5	0	2	normal
Allergic	17	3	0	1	normal
Kidney	9	2	0	0	normal
Other causes	54	20	0	0	normal
Abnormal gross bleeding (epistaxis, hematuria, melena, menorrhagic, postoperative, and retinal bleeding)	247	102	27	6	normal
Congenital thrombasthenia	3	3	3	0	normal
Miscellaneous (leukemia, enlarged spleen, etc.)	52	24	7	3	normal

Peck, Rosenthal, and Erf[3] have studied experimentally the production of purpura. With suction on a normal skin, the purpura is due to diapedesis of red cells through the intact capillary wall. With a toxin such as snake venom, there was a definite damage to the capillary wall, with rupture in some instances. In chronic thrombocytopenic purpura, there was often rupture of the vessels with suction alone, indicating a greater fragility of the capillary walls in clinical purpura.

We have summarized, in TABLE 2, the findings with reference to coagulation time, bleeding time, and platelet count in the 407 cases of abnormal hemorrhage without sufficient decrease in platelets or increase in coagulation time to cause the abnormal bleeding. While there is no significant change in these factors, such changes as do occur cannot be disregarded. It is very common to find some prolongation of coagulation time and/or bleeding time. It seems most probable that the factors which influence vascular permeability at the same time also influence the clotting factors. The change in clotting factors and platelets is secondary, however, to other changes.

Several case histories may be cited to illustrate this point.

Case 1. A woman, aged 27, was first seen for menorrhagia which necessitated transfusions to recover from the blood loss. A careful clinical study revealed no local or endocrine cause for the bleeding.

The blood study showed: red cells 4.75 millions, hemoglobin 78 per cent (12 g.), platelets 235,000, coagulation time fifteen minutes, bleeding time over ten minutes, clot retraction normal. The findings were the same in a second examination two months later.

It is apparent that the slight variation from normal in platelets, clotting and bleeding times is not sufficient to account for the abnormal hemorrhage. The increased coagulation time and bleeding time may, however, be an added factor.

Case 2. A girl, aged 17, had had menorrhagia since onset of menses. No local or endocrine cause for the bleeding could be demonstrated. The blood count showed 2.63 millions red cells, 44 per cent (6.8 g.) hemoglobin, and normal leukocytes. The coagulation and bleeding times were both twelve minutes. The platelets were 280,000 and the clot retraction was normal.

Here, again, the blood findings were insufficient to explain the abnormal bleeding, although they constituted an added influencing factor.

Case 3. A young woman, aged 23, had begun to have severe nosebleeds at the age of 13. She always bruised easily and on examination had a large ecchymosis of the right leg without apparent injury. She had rheumatic heart disease. There was no local cause for the nasal bleeding on examination of the nose. The tourniquet (Leed-Rumpel) test was positive. The coagulation time was eighteen minutes. The platelet count, coagulation time, and clot retraction were normal. There was no anemia or abnormality of the white cells.

Case 4. A 47-year-old woman had severe diabetes which was difficult to control. She also had a severe hemorrhagic retinitis. There was no other abnormal bleeding, and no anemia. The coagulation time on one occasion was twenty minutes and on another twenty-two minutes. The platelet count was 240,000, the clot retraction normal, and bleeding time one and a half minutes.

It seems most likely that the abnormal bleeding here was of nutritional origin and largely due to capillary endothelial injury, although the coagulation time was abnormal.

Many patients with abnormal hemorrhage not adequately explained by the changes in platelet count or clotting time, must have qualitative changes in the platelets and variation from normal in the blood proteins concerned in coagulation. Thus, one of us (Schneider) has found quite regularly an increased beta-globulin content and decreased albumin in the blood plasma of patients with hemorrhage retinitis such as case 4 cited above. The abnormal hemorrhage seen in pseudohemophilia or congenital thrombasthenia is dependent on quantitative changes in platelets. A typical example of this condition is the following:

Case 5. A boy who was six years old when first seen at Cleveland Clinic had always bled excessively from any injury. Also, he had always bruised easily. Three weeks before admission, he began to have bleeding from the throat after an attack of tonsillitis. There were a few subcutaneous hemorrhages on examination. The tourniquet test was positive. The symptoms cleared up temporarily with the use of snake venom. This boy has been under observation for seven years. With any injury, he bleeds excessively, requiring hospitalization.

Blood examination constantly shows a high platelet count (highest 720,000) and normal clot retraction. The coagulation time is always at the upper limit of normal or slightly prolonged. The bleeding time is always prolonged, at times as long as one and a half hours.

The child's mother has similar blood findings. She had severe nose bleeds at the age of 9 and missed school for one year. There was occasional menorrhagia, though never significant. Following delivery of her only child (case 5), hysterectomy was necessary to stop the uterine bleeding. She does not bruise easily and has no abnormal bleeding, although her bleeding time, when last tested, was one hour and forty-five minutes.

Many examples can be cited of purpura associated with rheumatic disease in which the purpura is very marked and often occurs with active arthritis. Purpura from drugs is frequently encountered clinically. Typically, these patients show no alteration in platelets or clotting factors, although not infrequently a slight reduction in platelets or some prolongation of bleeding and/or clotting time is encountered.

Discussion. It is apparent, from an analysis such as this, that most cases of clinical hemorrhagic diseases are not due to thrombocytopenia or to a disturbance in the clotting mechanism. The excessive bleeding of over three-fourths of the patients was due to increased capillary permeability.

With any patient with abnormal bleeding, the first procedure is to study the blood for hemorrhagic disease. If no blood dyscrasia is found, further search must be made for other factors causing increased capillary permeability. The more important of these are: (1) infection; (2) drugs; (3) toxemia; (4) allergy; and (5) disturbed nutrition. A Leede-Rumpel test should always be included in the clinical examination. If the test is positive, the capillary permeability is always impaired, although a positive test is also characteristic of thrombocytopenia.

Slight changes in platelets and clotting time which, by themselves, would not produce the bleeding of hemophilia or essential thrombocytopenic purpura, may well be a factor in abnormal bleeding due primarily to increased capillary permeability. The alteration in platelets and clotting are probably brought about by the toxic agents affecting the endothelial cells and intercellular cement.

The treatment of hemorrhagic disease not due to thrombocytopenia and abnormal clotting depends, primarily, on removing the offending agent, such as a drug or correcting a primary condition, as allergy. If the hemorrhagic disease is serious, transfusion may be of great value. Injections, intravenously, of calcium chloride or gluconate are often helpful. Many other agents, such as snake venoms, have been suggested. These are not of striking value.

CONCLUSIONS

Increased capillary permeability is by far the most frequent cause of clinical hemorrhagic disease.

In the series of 527 patients reported here, only 120 patients showed enough change in platelets or clotting factors to account for the disease.

Increased capillary permeability is due to changes in the intercellular cement substance or to primary damage to the endothelial cells.

Capillary damage may be due to infection, drugs, toxemia, allergy, or nutritional disturbances.

In abnormal bleeding resulting from increased capillary permeability, there may be a slight increase in bleeding and/or coagulation time, and reduction in platelets, insufficient in themselves to produce the abnormal bleeding. These are only added causative factors in bleeding primarily due to endothelial damage.

The treatment of this type of abnormal bleeding depends on finding and removing or treating the primary cause. Intravenous injections of calcium salts may help.

REFERENCES

1. **Quick, A. J.**
   1942. The Hemorrhagic Diseases and the Physiology of Hemostasis: 306. C. C. Thomas. Springfield, Ill.
2. **Zweifach, B. W.**
   1940. The structural basis of permeability and other functions of blood. Cold Spring Harbor Symp. Quant. Biol. **8**: 216-224.
3. **Peck, S. M., N. Rosenthal, & L. A. Erf**
   1937. Purpura, classification and treatment, with special reference to treatment with snake venom. Arch. Dermat. Syph. **35**: 831-867.

CLINICAL ASPECTS OF HYPOPROTHROMBINEMIA

By Charles S. Davidson and Henry J. Tagnon
Thorndike Memorial Laboratory, Second and Fourth Medical Services (Harvard), Boston City Hospital; and Department of Medicine, Harvard Medical School, Boston, Massachusetts

With the unfolding of our knowledge of the physiology of blood prothrombin and its relation to vitamin K, a better understanding has arisen of the diseases associated with hypoprothrombinemia, their diagnosis, and their rational treatment.

The discovery, by Dam and Schonheyder,[1] in 1934, that vitamin K cured chicks of hypoprothrombinemia produced by feeding a diet inadequate in this vitamin, led to much of the present advance in our knowledge. With the availability of vitamin K preparations, and with general use of methods for the determination of prothrombin, a number of hemorrhagic diseases with previously obscure etiology have been included in the group of conditions exhibiting a deficiency of prothrombin in the blood.

Vitamin K is provided in the normal diet. Probably an equally important source is its synthesis by bacteria in the gastrointestinal tract. It is fat-soluble and, therefore, for absorption, bile must be present in the gastrointestinal tract. Prothrombin is produced entirely by the liver, and for its synthesis, vitamin K is necessary. It is not known whether vitamin K acts as a component of an enzyme system or actually is a part of the prothrombin molecule itself. It has also been suggested that vitamin K acts by inhibiting a liver cathepsin which destroys prothrombin. Prothrombin circulates in the blood either as a plasma protein or closely associated with the plasma protein, and acts as an essential component in normal blood coagulation.

Recent experiments in animals throw further light upon the physiology of vitamin K and support the evidence that the vitamin is produced in the gastrointestinal tract.[2] Rats maintained on a diet deficient only in vitamin K, to which the poorly absorbed sulfonamide sulfaguanidine was added, were found by Black *et al.*[3] to develop hypoprothrombinemia. The blood prothrombin concentration was returned to normal by vitamin K administration. Rats maintained on a diet deficient in vitamin K without added sulfonamide did not develop prothrombin deficiency. It is thought that the action of the sulfonamide was due at least in a large part to its depressing action on the intestinal bacteria. Hypoprothrombinemia developing in man from sulfonamide administration has not yet been shown to exist.

These many factors which influence the available production of vitamin K by the body and the subsequent formation of prothrombin by the liver, create a number of possibilities for the clinical occurrence of hypoprothrombinemia. Lack of vitamin K in the diet, non-availability from disturbances in absorption, and failure of prothrombin synthesis due to a damaged liver represent the most important causes of hypoprothrombinemia. Many of these physiologic disturbances occur in diseases and other conditions in man and will be discussed with the known physiology of prothrombin and vitamin K as a basis.

A deficiency of prothrombin activity in the blood impairs its coagulability. In the clinical conditions discussed below, exhibiting a prothrombin deficiency, the coagulation time of the blood is not usually prolonged unless the prothrombin concentration is very low, *i.e.*, less than 10 per cent of normal. In spite of a normal coagulation time, an hemorrhagic tendency may exist when hypoprothrombinemia is present, for measurement of the coagulation time is not a sensitive test for small changes in blood coagulability.

There are no specific changes in the other blood constituents associated with hypoprothrombinemia. Bleeding time is usually normal. Blood platelet counts are within the normal range, unless the particular illness is associated with thrombocytopenia. Thus, in advanced liver disease, there may be thrombocytopenia as well as prothrombin deficiency.

The estimation of vitamin K in blood is not yet on a practical basis.

General Clinical Manifestations of Prothrombin Deficiency. The hemorrhagic manifestations of hypoprothrombinemia are varied. Subcutaneous hematomata, bleeding from venepuncture, mucous membrane bleeding, or post-traumatic bleeding have been reported. Kark and Souter[4] have divided hypoprothrombinemia into two groups: latent, and spontaneous hemorrhagic hypoprothrombinemia. The former exhibits prothrombin concentrations of between 20 and 35 per cent, the latter below 20 per cent of normal. In the first group, the patients had bleeding episodes following trauma or operation, while in the second group, the trauma was often insignificant or undetermined. The presence of small hematomata forming around the site of the venepuncture was found consistently in patients with a prothrombin of less than 35 per cent of normal. The manifestations of hemorrhagic disease of the newborn will be given special consideration.

Occasionally, the only indication of hypoprothrombinemia may be excessive bleeding from an organic lesion. Thus, in liver disease with portal obstruction and hypoprothrombinemia, bleeding from hemorrhoids or esophageal varices may be much more profuse than is usually found when the prothrombin concentration of the blood is normal.

CLINICAL CONDITIONS ASSOCIATED WITH PROTHROMBIN DEFICIENCY

Diseases Associated with Deficient Absorption of Vitamin K: *Obstructive Jaundice.* When bile is excluded from the gastrointestinal tract, homogenization of fats does not occur, and absorption of fat and fat-soluble materials may be incomplete or absent. Under such circumstances, the fat-soluble vitamins, including vitamin K, are poorly absorbed, so that hypoprothrombinemia may result.

The regular occurrence of hypoprothrombinemia in patients with obstructive jaundice accounts, at least in part, for the high surgical mortality in this condition. Wound bleeding often occurs and may be quite severe. The simple measure of administering bile salts to such patients, by mouth, will again permit the absorption of vitamin K. The administration of vitamin K without bile salts, by mouth, is, of course, ineffective. The vitamin may, however, be administered parenterally with rapid return of the blood prothrombin concentration to normal.

Disease of the Gastrointestinal Tract. A marked increase in the flow of the gastrointestinal contents may lead to deficient intestinal absorption. It is equally true that the failure of absorption of fat by any cause will lead to vitamin K deficiency and consequent hypoprothrombinemia. Therefore, it is to be expected that a prothrombin concentration below normal may be found in such conditions as sprue, idiopathic steatorrhea (celic disease), and high intestinal fistulae. Usually, in these individuals, the prothrombin deficiency is not sufficiently pronounced to produce a tendency towards bleeding, although increased bleeding may occur occa-

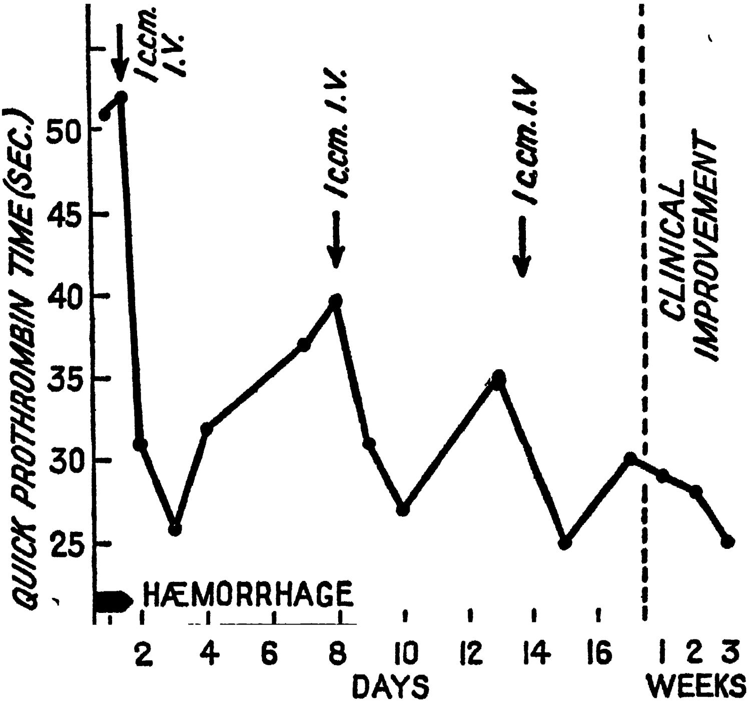


Figure 1. Effect of injections of vitamin K on the prothrombin time of a patient with ulcerative colitis. (From Kark & Souter[4]; reprinted by permission of *The Lancet.*)

sionally. A case illustrating this situation has been described by Kark, Souter, and Hayward.[5] Their patient had a very low prothrombin concentration, associated with poor fat absorption occasioned by idiopathic steatorrhea. Response to parenteral vitamin K was prompt, though not well maintained. A similar hypoprothrombinemia may occur in individuals with ulcerative colitis (FIGURE 1), regional ileitis, and chronic diarrhea of other causes. In the presence of any of these conditions, hypoprothrombinemia should be looked for, and if found, one would expect restoration of the blood prothrombin to normal to be effected through administration of vitamin K by the parenteral route.

Malnutrition. As has been stated, the discovery of hypoprothrombinemia associated with vitamin K deficiency was originally due to feeding chicks an inadequate diet. These diets were extracted in such a way that no fat-soluble vitamins were present. Therefore, it would seem possible that a true dietary deficiency of vitamin K could exist in humans. Kark and Lozner[6] have reported four cases of malnutrition with multiple vitamin deficiencies and mild prothrombin deficiency due to an insufficient supply of vitamin K. The blood prothrombin responded to the administration of the vitamin. All cases of pan-avitaminosis should be studied with respect to the presence of vitamin K deficiency and consequent reduction in the prothrombin level. More recently, Aggeler, Lucia, and Fishbon[7] reported a case of severe anorexia nervosa with marked hypoprothrombinemia. This responded promptly to the administration of adequate amounts of vitamin K.

Hemorrhagic Disease of the Newborn. Hemorrhagic disease of the newborn is a condition occurring in the first week of life, characterized by bleeding either internally or externally, and always associated with a diminished blood prothrombin concentration.

The hemorrhages may be from any site, quite often from the umbilical cord, gastrointestinal tract, vagina, or into the skin and subcutaneous tissues.

Although hypoprothrombinemia is a necessary criterion for the diagnosis of hemorrhagic disease of the newborn, all infants with hypoprothrombinemia do not bleed excessively. This is clear from the fact that most infants have a precipitous drop in prothrombin concentration on the day after birth. The low concentration continues for about three to six days and then rises to normal again, very rapidly. In contrast to this almost invariable drop in prothrombin concentration in the newborn, is the low incidence of hemorrhagic disease of the newborn. Thus, a lowered prothrombin concentration in itself is not sufficient for making a diagnosis. Hemorrhage in some form must be present. It is now well known that the hypoprothrombinemia is due to a deficiency of vitamin K, and that it may be prevented, or treated with the expectation of a prompt cessation of the bleeding when due to this cause, by the administration of vitamin K.

There are many early reports of bleeding in infants. In fact, as pointed out by Quick,[8] in 1940, Mosaic Law delays circumcision to the eighth day. The first clear description and documentation of the disease was made by Minot,[9] in 1852, under the title "On Hemorrhage from the Umbilicus in Newborn Infants, with an Analysis of Forty-six Cases". It was Townsend,[10] however, in 1894, who first called the condition "hemorrhagic disease of the newborn". Blood transfusion was then the best therapy. With the discovery of vitamin K by Dam and co-workers,[1] in 1934, and of its relation to blood prothrombin, the low prothrombin concentration in newborn infants pointed directly to the cause of the bleeding. Since that time, many reports have appeared, showing that vitamin K will prevent the fall of, or bring to normal, the prothrombin concentration of newborn infants. Whether there has been a reduction in the incidence of the disease by the routine use of vitamin K is still under discussion.[11, 12]

The clinical picture of the disease is easily recognized. The onset is in the first week of life and commonly on the second, third, or fourth days. After the twelfth day, the disease is seldom observed. The most characteristic feature is a spontaneous and persistent tendency to bleed, often multiple in origin. The bleeding has been observed in skin and subcutaneous tissue, from the nose and mouth, the conjuctivae, the mucous surface of the intestinal tract, intracranially, and into the serosa of the pericardium and the abdominal cavity. Certain organs sometimes show hemorrhages, as the kidneys, the adrenals, the thymus, the lungs, and the liver. Petechiae are rarely seen in hemorrhagic disease of the newborn.

The disease is self-limited with either death or spontaneous cure within the first week or at most 12 days of life. Without therapy, the mortality rate is high, being given variously as from 30 to 60 per cent.

Not all bleeding in the newborn is due to this disease. Trauma may be sufficient to cause ecchymosis, either due to a difficult delivery or a complicated postpartum course. It may, of course, be that such hemorrhagic episodes are much worse or continue longer with a low blood prothrombin concentration. The same holds true for bleeding from the umbilical cord stump. This occurrence is most commonly the result of insecure tying, but may be much more severe with a lowered prothrombin concentration.

Several other hemorrhagic diseases may occur in infants. Hemophilia rarely becomes manifest in the first year of life and is restricted to males; being a hereditary disease, it should not lead to confusion. Thrombocytopenic purpura occurs rarely in the newborn, but may be present in the newborn infant of a mother suffering from the idiopathic form of the disease. Vaginal bleeding *per se* often occurs in the newborn and is usually not a manifestation of hemorrhagic disease of the newborn, but is thought to be the result of an excess of estrogenic substances transmitted from the mother before birth.

Rarely, afibrinogenemia or fibrinogenopenia may be confused with hemorrhagic disease of the newborn. Simple examination of the blood clot should serve to differentiate these two, as the clot will be poor in quality or entirely absent in the diminution or lack of blood fibrinogen.

In the past, it was the custom often to attribute bleeding in the newborn to congenital syphilis or some other infection either hidden or patent. That these conditions may lead to bleeding is undoubted, but the differentiation from hemorrhagic disease of the newborn should be easily made.

The diagnosis of hemorrhagic disease of the newborn must be based on the clinical manifestations of abnormal bleeding from usually more than one site; on the low prothrombin concentration; and on the cessation of bleeding and rapid approach to normal prothrombin concentration after the administration of vitamin K.

The mechanism of vitamin K deficiency in the newborn is not well established. It is apparently not entirely due to a deficiency in the

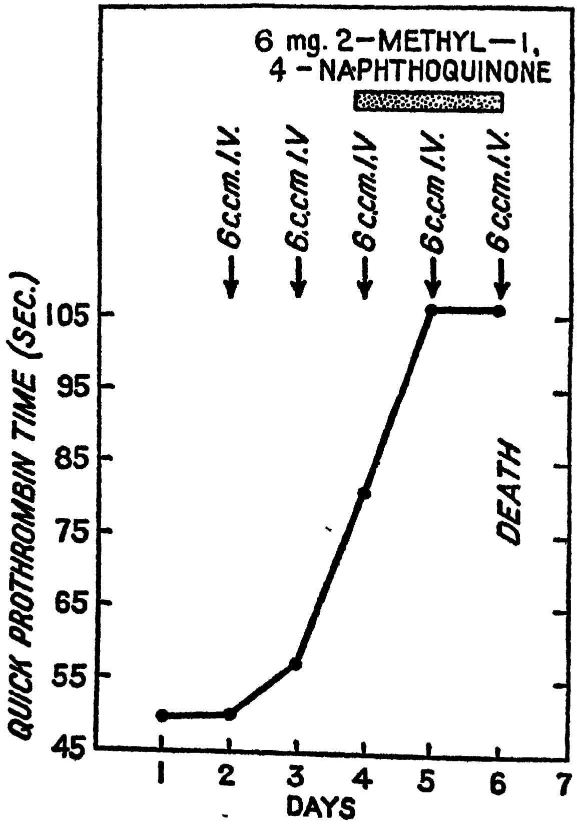


FIGURE 2. Lack of response to intravenous injections of vitamin K supplemented by oral administration, in a patient with acute yellow atrophy of the liver. (From Kark & Souter.[4] Reprinted by permission of *The Lancet*.)

mother, since her prothrombin concentration is usually within normal limits. Furthermore, the prothrombin concentration of the cord blood is usually within normal limits. It is well known that much of the vitamin K we absorb is synthesized by the intestinal flora, and it may be that the infant's relatively sterile gastrointestinal tract is unable to produce the vitamin.

Diseases Associated with Deficient Utilization of Vitamin K:

Liver Disease. Since the liver is the site of prothrombin formation, any severe hepatic disease which limits liver function may be associated with a blood prothrombin deficiency. Using phosphorus and chloroform to produce liver damage in a dog, Smith, Warner, and Brinkhaus[12] observed a profound fall in the concentration of prothrombin in the blood. Warner later found a similar result following partial hepatectomy in dogs. Clinically, a reduced prothrombin concentration in liver disease is often found. It is usually found in portal cirrhosis, but has been reported in primary and metastatic carcinoma of the liver, acute hepatitis, fatty infiltration, acute yellow atrophy, and catarrhal hepatitis. The absorption of vitamin K under these circumstances may be quite adequate, but the

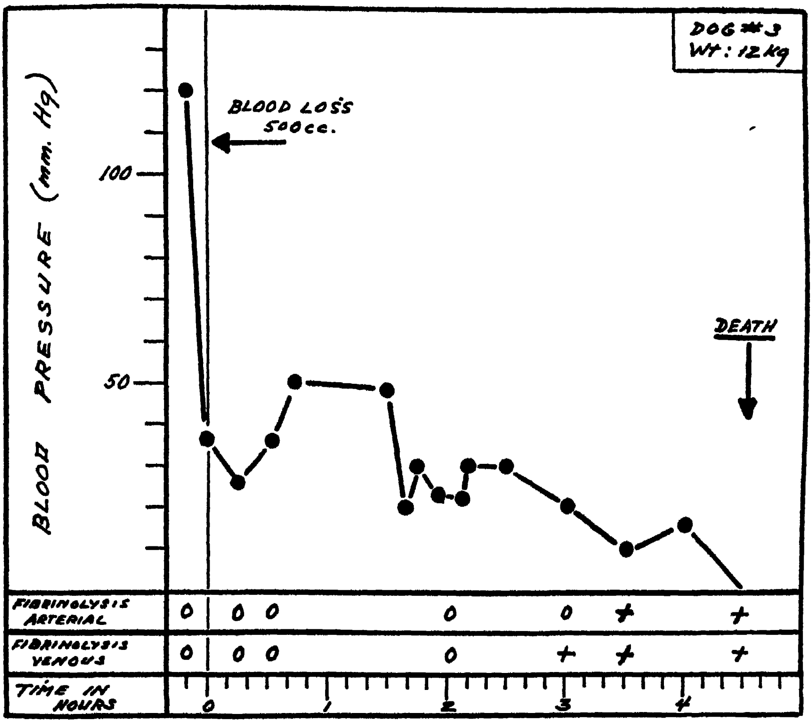


FIGURE 3. Comparison of fibrinolytic activity in venous and arterial blood in hemorrhagic shock in dog. (From Tagnon, Levenson, Davidson, & Taylor.[15] Reprinted by permission of the American Journal of the Medical Sciences.)

production of prothrombin by the injured or insufficient number of active liver cells is deficient. Therefore, the administration of vitamin K, orally or parenterally, will not increase prothrombin concentration in such individuals, unless there is a failure of absorption of vitamin K (FIGURE 2). In such instances, the administration of vitamin K may produce some rise in prothrombin concentration.

So frequently is hypoprothrombinemia associated with liver disease that the plasma prothrombin concentration and failure of adequate response to vitamin K administration have been proposed as an index of liver function. As in most liver function tests, considerable liver damage may be present without affecting the prothrombin level of the blood. In any event, there does not appear to be a direct relation between plasma prothrombin concentration and the other liver function tests usually employed. Lucia[13] has stated that the plasma prothrombin concentration cannot be depended upon as a good criterion of hepatic function.

Mechanism Unknown. *Shock.* It has been observed, in certain cases of traumatic shock in humans, that the prothrombin time was significantly prolonged. The same observation has been made in certain cases of hemorrhagic shock[14]. During shock produced by hemorrhage in dogs, the prothrombin time becomes progressively prolonged and the plasma fibrinogen decreases (TABLE 1, FIGURE 3). The mechanism of the production of hypoprothrombinemia in shock is not clear. Among the possible causative agents, one should consider either an acute vitamin K deficiency, since it has been shown that in a state of shock there is an increased demand for vitamins, or liver insufficiency caused by the anoxemia of shock. Another explanation could be that the prothrombin is destroyed by the proteolytic enzyme which is known to appear in the

TABLE 1*

PROTHROMBIN IN HEMORRHAGIC SHOCK (DOGS)
Quick Prothrombin Time (Seconds)

		Hour after beginning of shock					
Dog No.	Control	1	2	3	4	5	6
6†	13	14	17	20	Died		
7	11	12	19	24	32	Died	
8	11	12	Died				
9	11	11	12	—	—	13	17
10	12	15	19	Died			
11	12	14	17	19	23	Died	
12	13			20	26	Died	
13†	11	11	13	15	18	21	Died

* From Tagnon, Levenson, Davidson, & Taylor.[15] Reprinted by permission of The American Journal of the Medical Sciences.
† No anesthesia.

circulating blood in shock. At the present time and before more experimental evidence is obtained, one has to consider the hypoprothrombinemia of shock as still unexplained. It may, however, have practical significance in that it may provide the explanation for the persistence of the hemorrhage once it has started: a vicious circle becomes established, in virtue of which bleeding creates optimal conditions for more bleeding. It is probable that when the surgeon gets ready to operate on a patient in shock, he should pay attention not only to the restoration of the blood volume and the oxygen capacity of the blood, but also to the restoration of a normal prothrombin time. It is quite possible that this last requirement will be satisfied only if the fluids used for the restoration of the blood volume contain an adequate amount of prothrombin.

Drugs Affecting Blood Prothrombin Concentration:

Dicumarol. It was noted, some years ago, that cattle feeding in certain areas developed a marked bleeding tendency. This condition has been referred to as "spoiled sweet clover disease" in cattle. Link and his co-workers have isolated and synthesized the compound 3-3′-methylene-bis-(4-hydroxycoumarin) from spoiled sweet clover which, when in-

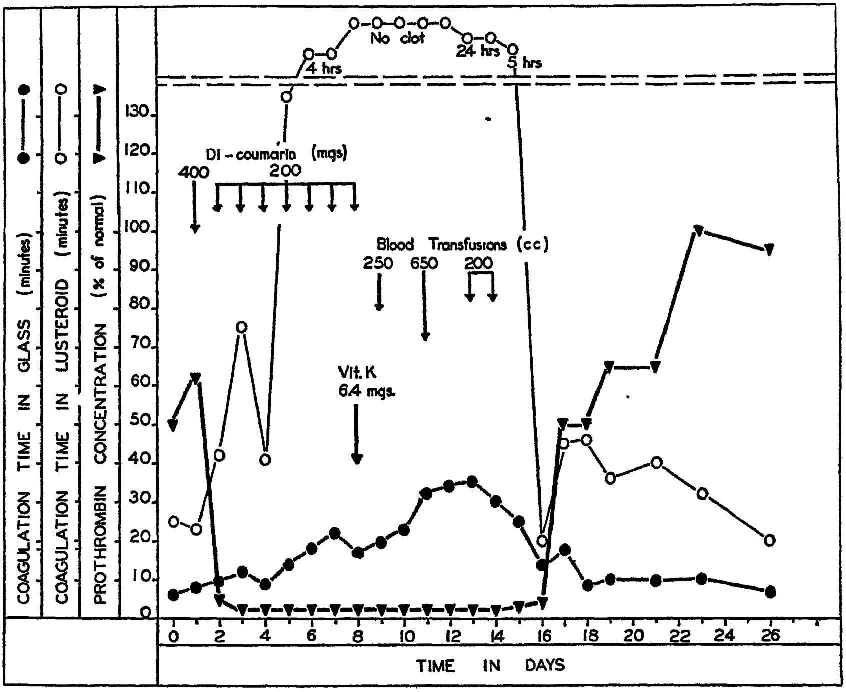


FIGURE 4. Changes in the coagulation time in glass and lusteroid and the prothrombin concentration following the administration of 1.8 g. of dicoumarin. Effect of the administration of vitamin K and whole blood transfusions. (From Davidson & MacDonald.[16] Reprinted by permission of The American Journal of the Medical Sciences.)

gested by experimental animals, resulted in the usual hemorrhagic symptoms of sweet clover disease. The hemorrhagic condition produced is due to a diminution of the blood prothrombin concentration. The use of the compound as a substitute for heparin has been proposed. Unlike heparin, the compound has no action *in vitro,* may be given by mouth and must be administered for twenty-four or forty-eight hours before its effect is noted. Often an even longer period of time must be allowed for the blood coagulability to return to normal after administration is stopped (FIGURE 4). An antagonistic effect has been shown to exist between 3-3′-methylene-bis-(4-hydroxycoumarin) and large amounts of vitamin K with respect to the synthesis of prothrombin (FIGURE 5).[15]

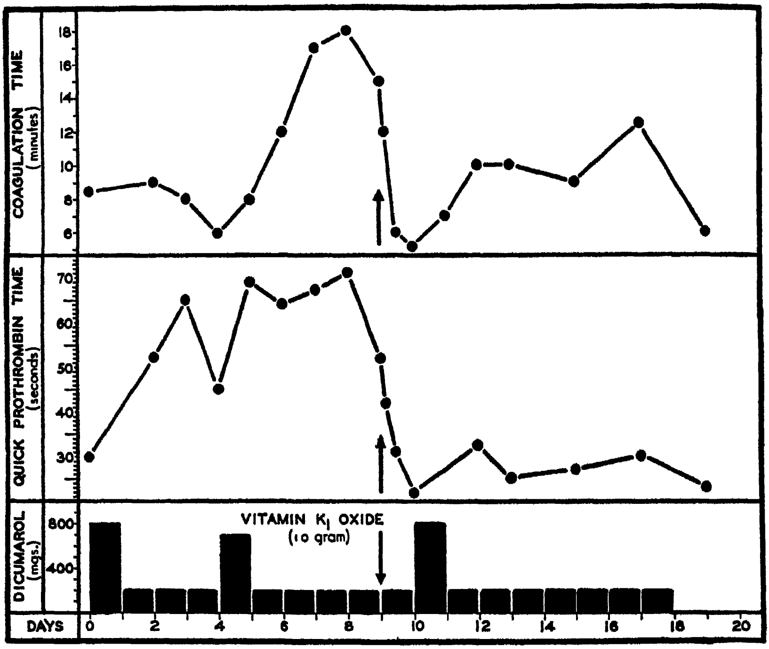


FIGURE 5. Antagonistic effect of vitamin K_1 oxide upon the anticoagulant effect of dicumarol in spite of continued administration of dicumarol. (From Davidson, Freed, & MacDonald. 1945. Reprinted by permission of The American Journal of the Medical Sciences (210: 684.))

Salicylates. The occasional appearance of hemorrhagic manifestations in salicylate poisoning has been noted for many years. An adequate explanation was not available until Link and co-workers[16] showed that salicylate administered to rats on a vitamin K deficient diet developed hypoprothrombinemia. These workers showed that when dicumarol was degraded *in vitro,* salicylic acid was formed. They suggested this as a possible explanation of the action of dicumarol *in vivo.* Both Shapiro and co-workers,[17] and Meyer and co-workers,[18] have extended Link's findings in rats to man and have demonstrated hypoprothrombinemia from the administration of salicylate even without restricting vitamin K in the diet. Thus, an explanation for the hemorrhagic phenomena in salicylism has been made, but whether Link's suggestion to explain the effect of dicumarol by degradation to salicylate in the body is correct, must await the demonstration of salicylate formed in the body during the administration of dicumarol.

TREATMENT

Vitamin K. Natural vitamin K exists in two forms: Vitamin K_1 and Vitamin K_2. Both are substituted naphthoquinones, each with one long side chain; vitamin K_1 being 2-methyl-3-phytyl-1, 4-naphthoquinone, and vitamin K_2 probably 2-methyl-3-difarnesyl-1, 4-naphthoquinone. Vitamin K_1 is derived from alfalfa and K_2 from putrefied fish meal. Vitamin K_2 has about 60 per cent of the activity of vitamin K_1. Vitamin K_1 has been synthesized, but not vitamin K_2. (Both are fat-soluble and, as with other fat-soluble materials, depend for their absorption upon the presence of bile salts in the gastrointestinal tract.) Not long after the identity of vitamin K_1 became known, the synthesis of other compounds having vitamin K-like activity began. Among these is the simpler compound, 2-methyl-1, 4-naphthoquinone, which, as Fieser[19] points out, is not synthetic vitamin K but a synthetic compound having vitamin K-like activity. Vitamin K_1 is apparently non-toxic even in large doses, whereas 2-methyl-1, 4-naphthoquinone, the most potent of the compounds, in animal experiments may be toxic in large doses. The preparations having vitamin K activity available for therapeutic use are chiefly vitamin K_1 for oral use, in the presence of bile salts, and water soluble compounds of 2-methyl-1, 4-naphthoquinone, which may be given intravenously or orally without bile salts. Vitamin K_1 may be prepared in emulsions for intravenous use, or may be given intramuscularly in oils. There is some evidence that vitamin K_1 has a longer duration of activity than the simpler naphthoquinones, though both become effective at approximately the same time after administration.

The dosage of the compound varies with its potency and according to the use for which it is desired. Using 2-methyl-1, 4-naphthoquinone as the standard, 5 to 10 mg. are usually sufficient to bring the prothrombin concentration to normal in simple vitamin K deficiency, as in obstructive jaundice. Less is required in hemorrhagic disease of the newborn. The action of the compound measured by a rising prothrombin concentration begins shortly after administration and reaches a peak in from 6 to 15 hours. Preparations given parenterally have a somewhat more rapid action.

Other Measures. Prothrombin in pure form is not available, but fresh whole blood and plasma contain it. Their administration may be of value when vitamin K cannot be expected to act, as in liver disease. In dicumarol poisoning, when the prothrombin concentration is very low, blood transfusion exerts little, if any, effect. Quick[20] has recently shown that prothrombin may be composed of calcium and two components called prothrombin A and B. Storage of plasma reduces only component A, whereas component B is reduced after the administration of dicumarol.

BIBLIOGRAPHY

1. **Dam, H.**
   1935. Nature **135**: 652.
2. **Greaves, J. D., & C. L. A. Schmidt**
   1937. Proc. Soc. Exp. Biol. & Med. **37**: 43.
3. **Black, S., R. Overman, C. H. Elvejem, & K. P. Link**
   1943. J. Biol. Chem. **145**: 137.
4. **Kark, R., & A. W. Souter**
   1940. Lancet **1**: 1149.
5. **Kark, R., A. W. Souter, & J. C. Hayward**
   1940. Quart. J. Med. **9**: 247.
6. **Kark, R., & E. L. Lozner**
   1939. Lancet **1**: 1162.
7. **Aggeler, P. M., S. P. Lucia, & H. M. Fishbon**
   1942. Am. J. Dig. Dis. **9**: 227.
8. **Quick, A. J.**
   1942. The Hemorrhagic Diseases and the Physiology of Hemostasis. Thomas. Springfield, Ill.
9. **Minot, F.**
   1852. Am. J. Med. Sci. **24**: 310.
10. **Townsend, C. W.**
    1894. Arch. Pediat. **11**: 559.
11. **Clifford, Stewart H.**
    1939. J. Pediat. **14**: 333.
12. **Sanford, H. N., L. Shmigelsky, & J. M. Chapin**
    J. A. M. A. **118**: 697.
13. **Smith, H. P., E. D. Warner, & K. M. Brinkhous**
    1937. J. Exp. Med. **66**: 801.
14. **Lucia, S. P., & P. Aggeler**
    1941. Am. J. Med. Sci. **201**: 326.
15. **Tagnon, H. J., S. M. Levenson, C. S. Davidson, & F. H. L. Taylor**
    1946. Am. J. Med. Sci. **211**: 88.
16. **Davidson, C. S., & H. MacDonald**
    1943. N. Eng. J. Med. **229**: 353.
17. **Link, K. P., R. S. Overman, W. R. Sullivan, C. F. Huebner, & L. D. Scheel**
    1943. J. Biol. Chem. **147**: 463.
18. **Shapiro, S., M. Redish, & H. A. Campbell**
    1943. Proc. Soc. Exp. Biol. & Med. **53**: 251.
19. **Meyer, D. D., & H. Beryl**
    1943. Proc. Soc. Exp. Biol. & Med. **53**: 234.
20. **Fieser, L. F.**
    1941. Ann. Int. Med. **15**: 648.
21. **Quick, A. J.**
    1943. Am. J. Physiol. **140**: 212.

DISCUSSION OF THE PAPER

Dr. L. B. Jaques:

While I agree with Dr. Davidson that the tendency is to use the one-stage method only for the determination of prothrombin, I would suggest that the two-stage method can provide data not obtainable by the one-stage method. Besides, when a number of samples are to be tested, the two-stage method as conducted by Herbert (Biochem. J. **34**:1554.) should not be too cumbersome for routine investigations.

I was particularly interested in the graph showing the intravenous administration of protamine. In our original communication, we counseled caution in this use of protamine, due to the marked toxicity of protamine in the dog. Recent studies by Dr. Fidlar and myself have suggested that this toxicity is an anaphylactoid response of moderate severity, but limited to the canine species, and that the toxicity of protamine may not be significant clinically. I would like to ask Dr. Davidson about symptoms following the injection of protamine and, also, how many patients have received protamine.

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Volume XLIX, Art. 5. Pages 661-866
September 24, 1948

RECENT STUDIES IN THE MECHANISMS OF EMBRYONIC DEVELOPMENT*

Consulting Editor: ROBERTS RUGH

CONTENTS

PAGE
Oöplasmic Segregation in Relation to Differentiation. BY DONALD PAUL COSTELLO 663
The Organization of the Amphibian Egg during Fertilization and Cleavage. BY G. FANKHAUSER 684
Significance of the Cell Membrane in Embryonic Processes. BY JOHANNES HOLTFRETER 709
Prospective Areas and Differentiation Potencies in the Chick Blastoderm. BY DOROTHEA RUDNICK 761
Biochemical Differentiation during Amphibian Development. BY E. J. BOELL 773
Form Changes during Pre-Gastrular Development. BY J. S. NICHOLAS 801
The Role of Nerves in Amphibian Limb Regeneration. BY S. MERYL ROSE 818
Quantitative Studies on Locomotor Responses in *Amblystoma* Larvae Following Surgical Alterations in the Nervous System. BY S. R. DETWILER 834
Functional Polarization in Developing and Regenerating Retinae of Transplanted Eyes. BY L. S. STONE 856

---

* This series of papers is the result of a Conference on The Mechanics of Development, held by the Section of Biology of The New York Academy of Sciences on January 10 and 11, 1947, with Roberts Rugh of New York University as Organizing Chairman.

Publication made possible through a contribution from the General Funds of The New York Academy of Sciences.

Copyright, 1948, by The New York Academy of Sciences

THE MERRYMOUNT PRESS, BOSTON, U. S. A.

OOPLASMIC SEGREGATION IN RELATION TO DIFFERENTIATION

By DONALD PAUL COSTELLO

University of North Carolina, Chapel Hill, North Carolina, and The Marine Biological Laboratory, Woods Hole, Massachusetts

OF the many processes which take place during the maturation, fertilization, and early cleavage of invertebrate eggs, it is apparent that some are more significant than others in causally contributing to differentiation. The fact of parthenogenesis indicates that the contributions of the spermatozoon to development, particularly its nuclear contribution, cannot be of primary significance. The fact that certain ova, under experimental conditions, may show various degrees of differentiation without cleavage, indicates that neither the mitotic mechanism, nor cleavage, nor the cleavage pattern, nor cell boundaries are of essential importance. It is likewise clear, from experimental evidence as well as by genetic assumption, that the differential competence of the different blastomeres of the early embryo cannot be accounted for on the basis of nuclear, chromosomal, or genic differences between the blastomeres. We are, thus, led to the conclusion that the factors necessary for differentiation are those producing the localization of the cytoplasmic areas of specific potency—or, if one wishes to use another terminology, the factors producing cytoplasmic fields and gradients. The study of such factors may be approached from the standpoint of metabolic measurements, in terms of differential metabolic rates and enzyme distribution. However, I believe that a more direct *biological* approach to the problem is afforded by the study of the process that has been termed oöplasmic segregation.

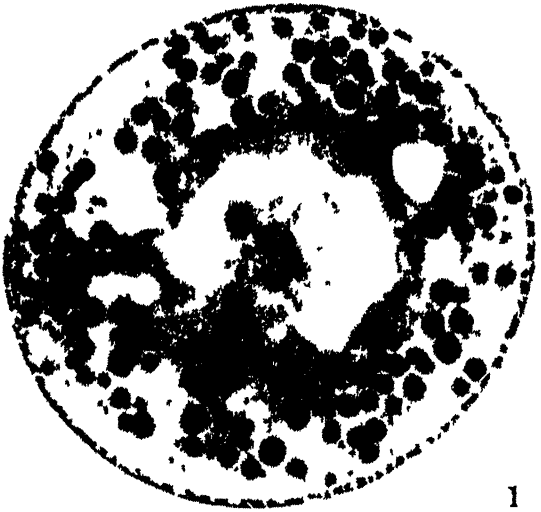
The segregation of the visible cytoplasmic elements in the eggs of marine invertebrates was early described by Wilson (1892), Conklin (1905), Lillie (1906), and others, and has been more recently studied by Spek (1930, 1934a, 1938) under the term, *bipolar differentiation*. It is the purpose of the present paper to review the recorded facts in relation to oöplasmic segregation and to make some suggestions concerning the significance of the process.

The phenomenon of visible oöplasmic segregation is particularly striking in eggs showing the so-called "determinate" type of cleavage, and may be initiated, in different forms, at the time of, or prior to, germinal vesicle breakdown, during polar body formation, or at fertilization or parthenogenetic activation. Invertebrate eggs of the various animal groups may be divided into different categories as regards the time of onset, and the pattern, of oöplasmic segregation. There are also different relationships between the pattern of segregation and the cleavage pattern. Because of these various types of segregation patterns, Spek's term, bipolar differentiation, is a misnomer. Since there is an original polarity of the egg that can be traced back to the earliest stages of the oögonium, it is recognized that the simplest type of oöplasmic segregation is essentially a re-polarization of certain of the egg constituents.

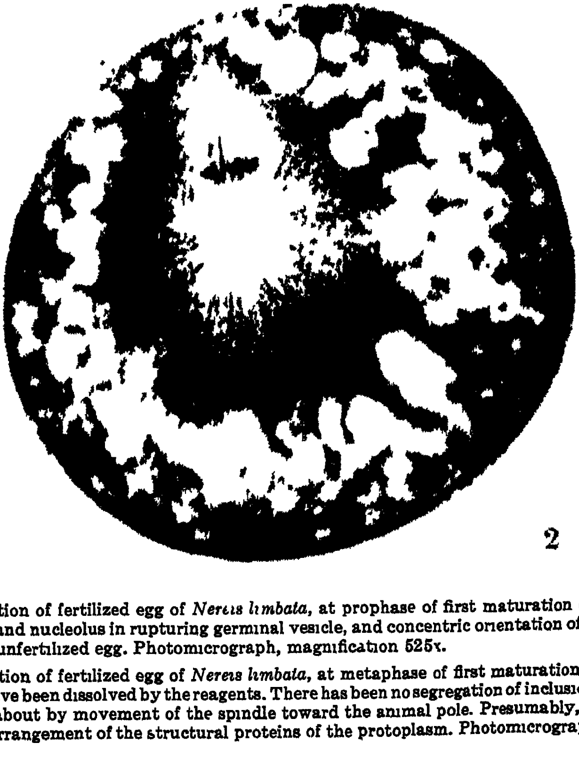
Oöplasmic Segregation in the Egg of NEREIS. The bipolar pattern of oöplasmic segregation is, perhaps, best exemplified by the *Nereis* egg. Spek (1930) studied the process in the egg of *Nereis dumerilii*, using vital staining methods. The process of segregation of the oöplasmic constituents does not begin until after extrusion of the polar bodies (FIGURES 1 and 2). It is especially apparent at the four-cell stage, as figured by Spek (1934a) for *Nereis limbata*, where the animal hemisphere gives an alkaline reaction with indicator dyes and the vegetal hemisphere gives an acid one. This has been discussed in considerable detail by Costello (1945a) for the egg of *Nereis limbata*. Since the animal hemisphere is destined to give rise in development to the ectodermal quartets of micromeres, whereas the vegetal hemisphere produces the endodermal macromeres, the implication of Spek's work is that there is a causal relationship between this "bipolar segregation" of acid and alkaline "protoplasm" and the differentiation of endoderm and ectoderm. I shall return later to the discussion of this point.

In normal eggs of *Nereis*, the time of onset of visible oöplasmic segregation is shortly after the final incorporation of the sperm head into the egg. Since an accentuated aggregation of cytoplasmic components is obtained in *Nereis* eggs from which the activating spermatozoon is removed (along with the vitelline membrane) by alkaline sodium chloride (Costello, 1945b), sperm entrance is not a necessary prerequisite. Certain salt solutions (Spek, 1930, 1934b) may also induce segregation in unfertilized eggs.

Oöplasmic Segregation in the Egg of CHAETOPTERUS. In the egg of *Chaetopterus*, segregation takes place much earlier than in that of *Nereis*, and it is essentially completed by the metaphase of the first maturation division (Lillie, 1906). The ectoplasm of the ovarian egg covers the free hemisphere and ends a short distance below the equator, so that the endoplasm comes to the surface in the vegetal hemisphere. There is usually also a small ectoplasmic defect at the animal pole, where the endoplasm comes to the surface. After the egg is shed by the female *Chaetopterus* into sea water, the germinal vesicle ruptures and a series of movements of egg substance takes place. The ectoplasm flows toward the vegetal pole, covering the exposed endoplasm. The original polar defect enlarges and the maturation spindle becomes attached there. There is also a redistribution of the different types of endoplasmic granules *a*, *b*, and *c* (see Lillie, 1906, Figures 1, 2, 5, and 25). This distribution of substances corresponds, in many respects, to the future embryonic areas and is maintained from this time on. At the first cleavage, the polar



1

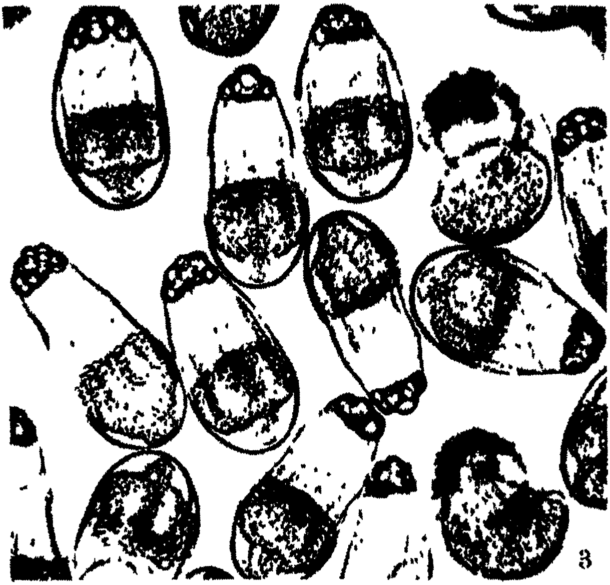


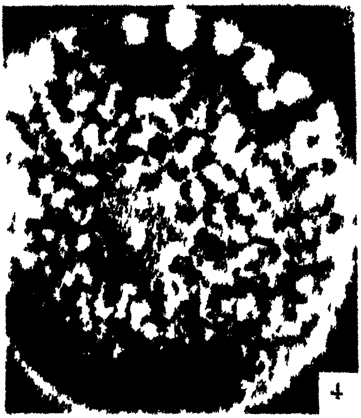
2

FIGURE 1. Section of fertilized egg of *Nereis limbata*, at prophase of first maturation division. Note chromosomes and nucleolus in rupturing germinal vesicle, and concentric orientation of ooplasmic inclusions as in unfertilized egg. Photomicrograph, magnification 525×.

FIGURE 2. Section of fertilized egg of *Nereis limbata*, at metaphase of first maturation division. The oil and yolk have been dissolved by the reagents. There has been no segregation of inclusions other than that brought about by movement of the spindle toward the animal pole. Presumably, the spindle is formed by rearrangement of the structural proteins of the protoplasm. Photomicrograph, magnification 590×.

bodies and ectoplasmic defect are transmitted to the *CD* blastomere, and at the third cleavage the polar defect is found on the 1*d* blastomere, which gives rise to the apical tuft. A more detailed account of the be-





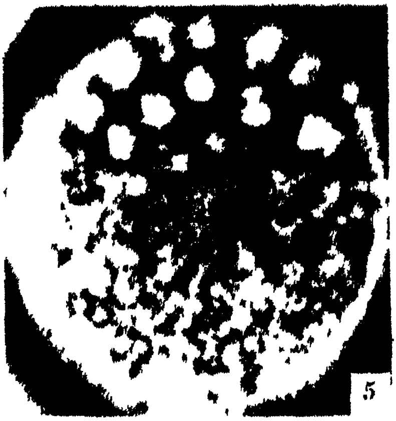


FIGURE 3. Unfertilized *Nereis* eggs centrifuged 60 minutes at about 6000 times gravity, photographed 5 minutes after centrifuging. The large oil droplets mark the centripetal end. All inclusions are sedimented into strata. Magnification 135×.

FIGURE 4. *Nereis* egg centrifuged 24 minutes in 0.73 M sucrose at 6000 times gravity, inseminated immediately, and photographed 53 minutes later. Note oil droplets marking centripetal pole. Polar area, in surface view, is at equator. Magnification 410×.

FIGURE 5. *Nereis* egg centrifuged 10 minutes in 0.73 M sucrose at 6000 times gravity, inseminated immediately, and photographed 70 minutes later. Note oil droplets, marking centripetal pole, at top, polar area and first polar body in surface view at equator. Magnification 410×.

havior of the *Chaetopterus* egg during this segregation period is given by Lillie (1906) and summarized with special reference to the possible role of the polar defect, by Costello (1945a).

The Mechanism of Oöplasmic Segregation. With these brief descriptions of two examples of the process of oöplasmic segregation, we have materials for a consideration of the mechanism involved. Spek (1930, 1934a, 1938) speaks of the process as being brought about by a "self-cataphoresis," but there is no evidence for a flow of electric current through the cell, and no evidence that such a distribution of substances is obtained when a potential difference is superimposed upon the cell from outside. Nor will the action of gravity (or magnified gravity, as in a centrifuge) in sedimenting the contained oöplasmic inclusions produce a similar pattern. For example, in oöplasmic segregation in the *Nereis* egg, oil and yolk go toward the vegetal pole, whereas in the centrifuge (FIGURE 3), oil moves centripetally, yolk centrifugally, to opposite directions, which are usually at right angles to the animal-vegetal axis (FIGURES 4 and 5). This is due to orientation of the somewhat disc-shaped egg in the centrifuge, with polar axis (short axis of the disc) at right angles to the direction of the centrifugal force.

An explanation of oöplasmic segregation was recently proposed (Costello, 1945a) in terms of the Teorell "diffusion effect." As illustrated in FIGURE 6, the Teorell system is as follows: Across a permeable bound-

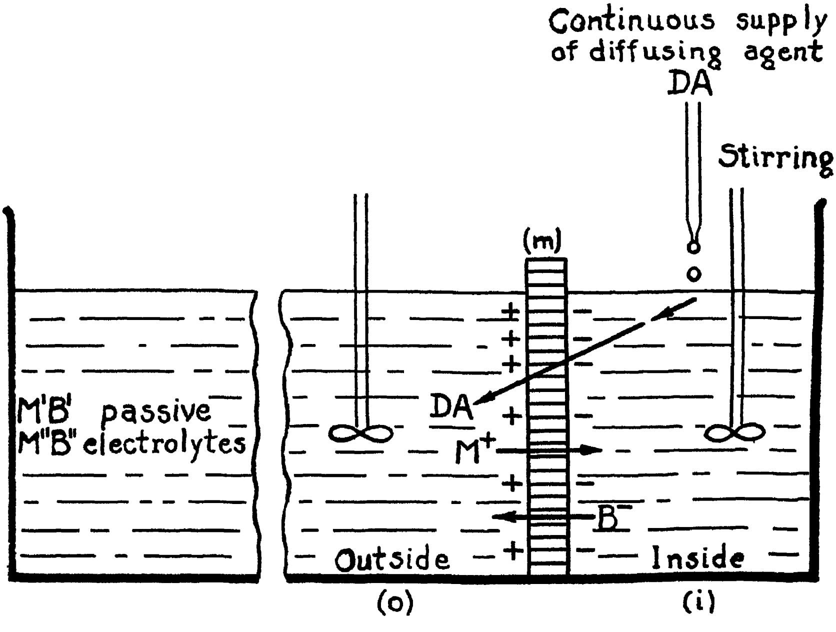


FIGURE 6. Diagram to illustrate Teorell's experimental arrangement for demonstration of the "diffusion effect." (After TEORELL, 1937.)

ary (*m*) there is assumed to be present a constant difference in concentration of either the cation D^+ or the anion A^-. The maintenance of this condition is accomplished by the continuous addition of the substance DA to the small volume (*i*) [= inside], causing DA to act as a "diffusion agent" which steadily diffuses across (*m*) into the part (*o*) [= outside]. In (*o*), a fixed, constant composition is maintained by keeping the volume large.

The continuous steady diffusion of DA was shown by Teorell (1935a, 1937) to influence the distribution of other electrolytes present, denoted by M′B′, M″B″, etc., which were not participating in any active diffusion and were therefore called "passive ions." The membrane is permeable to these passive ions. It was assumed that the D^+ ions had a higher mobility in the boundary (*m*) than the A^- ions, producing an electrical potential across the boundary. Starting with the initial state of equal concentrations of M^+ and of B^- on the two sides of the membrane (*m*), the electrical potential causes an inward migration of M^+ and an outward migration of B^-. Finally the concentration gradients become sufficiently large to balance the electrical gradient and the system approaches a steady state. Thus, the M^+ ions accumulate and B^- ions decrease in amount inside. Therefore, a diffusion of one electrolyte may produce, inside, an accumulation or impoverishment of other cations or anions, depending upon the mobilities of the ions of the diffusing agent. This effect upon ionic distribution is called the diffusion effect.

Teorell's concept of diffusion effect was later extended (1935b) to apply to any electrically charged particle, regardless of size. The diffusion potentials present in the system would thus move positively charged particles into the negative part of the diffusion potential field, and negatively charged particles in the opposite direction. It was emphasized that this "diffusion effect" upon ionic and colloidal distribution is not a cataphoretic effect, because no current is flowing and no external E.M.F. is applied. It is brought about by exchange of charged particles due to differences in mobility of the ions of the diffusing substance within the membrane.* Teorell (1935c) also devised a scheme for studying ionic distribution within a thick diffusion layer—the so-called multi membrane arrangement.

In my 1945 paper, I proposed a biological analogue of the Teorell scheme to account for oöplasmic segregation (FIGURE 7). If there were a continuous supply of diffusing substance entering the cell across a special area of the cell membrane, such as the polar area, diffusion potentials might be set up within the egg, with part of the protoplasm (such as the protein framework) acting as a multimembrane. If electrically charged particles, regardless of their size, were subjected to the influence of these diffusion potentials, positively charged particles would be moved toward one pole of the egg, negatively charged particles toward the other. In this earlier paper, I considered at some length the assumptions basic to the application of the Teorell scheme in such a case, and the reader is referred there for further details. I should like to point out that Harrison (1945) has recently visualized the ovum (Harrison's Text figure 12) in a manner essentially corresponding to my last figure, with a protein framework and distribution within this framework of two types of substances. Harrison distinguishes two types of polarity in the ovum: a polarity of direction (Richtungspolarität) and a polarity of stratification (Schichtungspolarität). The former is equivalent to what I have called the original polarity of the egg, the latter to what is brought about by the process of segregation.

---

* The normal difference in mobility of the ions of a substance is sufficient, but this mobility difference may be accentuated by certain types of membranes (Teorell, 1937).

If oöplasmic segregation is initiated when the egg enters sea water, or when some event, such as fertilization, alters the membrane permeability to permit entry of the diffusing substance, then the diffusing substance is probably either one of the common constituents of sea water, or a constituent of the outer surface of the egg which is able to enter the egg after its permeability is altered. Spek (1930, 1934b) has induced an artificial "bipolar differentiation" in the eggs of *Nereis* and *Asterias* by means of potassium chloride. Mead (1898) induced *Chaetopterus* eggs to complete their maturation by the addition of a small quantity of potassium chloride to sea water. It is also this substance which induces *Chaetopterus* eggs to undergo differentiation without cleavage (Lillie, 1902, 1906). There is, therefore, the possibility that the diffusing substance may be potassium chloride. However, it is very important

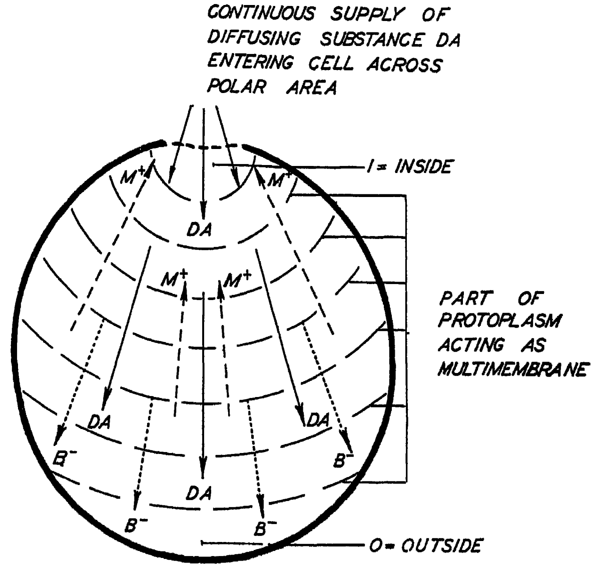
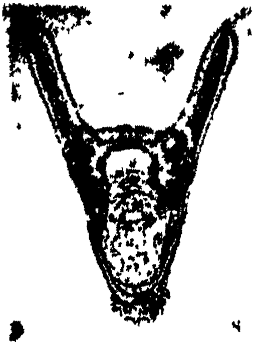
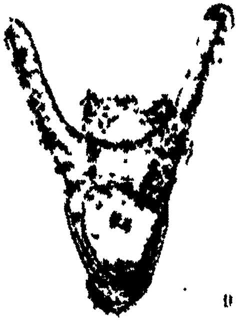

CONTINUOUS SUPPLY OF DIFFUSING SUBSTANCE DA ENTERING CELL ACROSS POLAR AREA
I = INSIDE
M+
M+
DA
M+
M+
PART OF PROTOPLASM ACTING AS MULTIMEMBRANE
DA
DA
B-
B-
DA
B-
B-
O = OUTSIDE

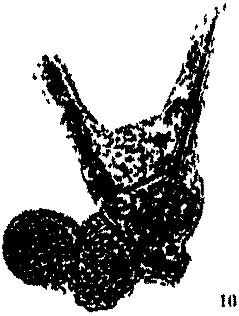
FIGURE 7. Biological analogue of Teorell scheme, employing "multimembrane" or thick diffusion layer to bring about segregation of particles M^+, B^- within cell. (After COSTELLO, 1945a.)

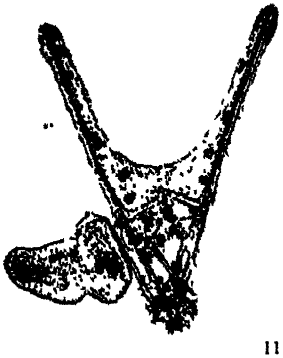
that the diffusing substance should show a considerable difference in mobility of its ions within the "membrane" [*i.e.*, (*m*)], in order to set up a diffusion potential of any great magnitude. The mobilities of the potassium and of the chloride ion are almost the same in simple solutions. We do not know with certainty what their relative mobilities would be within the egg protoplasm.

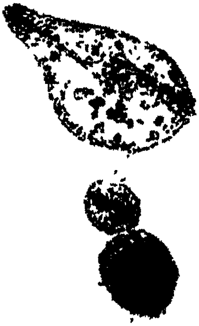
It is a well-known fact that modification of the external medium, as by the addition of a simple salt such as lithium chloride, leads to modification of development. Some examples of lithium-induced exogastrulation in the sea urchin, *Arbacia punctulata*, are shown in FIGURES 8–13.*


8


9


10


11


12

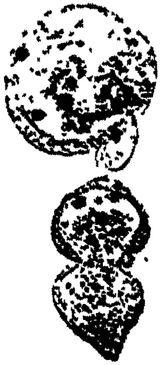

13

FIGURE 8. Normal, living pluteus of *Arbacia punctulata* about 40 hours old. Viewed from aboral surface, magnification 150x.

FIGURE 9. Same, viewed in median optical section.

FIGURE 10. Exogastrulated pluteus, 48 hours old, developed from *Arbacia* egg treated for 4 hours with mixture of 0.54 M LiCl (20 parts) and sea-water (80 parts), beginning 53 minutes after insemination (during 2-cell stage). Magnification 150x.

FIGURE 11. Exogastrulated pluteus, 52 hours old, same treatment as FIGURE 10. Magnification 130x.

FIGURE 12. Exogastrulated pluteus, 57 hours old, same treatment and magnification as FIGURE 10. Effect more extreme.

FIGURE 13. Exogastrulated pluteus, 52 hours old, same treatment and magnification as FIGURE 12. Effect still more extreme.

---

* Observations made by Mrs. Marjorie Hopkins Fox and the author in 1945.

It is possible that the lithium acts through an effect on the diffusion potential and oöplasmic segregation. *Arbacia* embryos from a given culture which has been subjected to a specific dosage of lithium chloride to induce exogastrulation, exhibit a remarkable variation in the degree of developmental modification. Child (1940) has exhaustively discussed this variation in susceptibility to lithium. He suggests that aggregation at the bottom of the container produces environmental conditions different from those surrounding isolated individuals. One might argue that an internal protoplasmic gradient would tend to produce just such a high degree of variability if the eggs come to rest with their polar axes oriented at any angle with respect to the bottom of the dish. That is, in some cases, the external chemical gradient effect would be added to the original polar gradient; in other cases, the two would act in opposite directions; and in still other cases, at all possible angles to each other.

Since the early experiments of Herbst (1892), it has been recognized that lithium has at least two distinct effects on echinoderm eggs. It produces exogastrulation (*i.e.*, separation of endodermal from ectodermal and mesodermal structures with essentially complete differentiation of all three) and inhibition of development, which may produce either inhibitory ectodermal or endodermal modifications. Perhaps a third effect is endodermalization of prospective ectoderm. If we can attribute these separate effects to different physico-chemical causes, I should postulate that the primary exogastrulation is brought about by a physical effect of the lithium ion through some such mechanism as the Teorell diffusion effect acting on oöplasmic segregation, accentuating the separation of ectodermal and endodermal factors responsible for differentiation. The secondary inhibition of development could be assumed to be due to an inhibitory effect of lithium on certain enzyme systems (however, contrast Pease, 1942).

There are a number of unexplained effects of certain other external agents in embryonic development. As a working hypothesis, it might be possible to invoke the Teorell diffusion effect as the causative mechanism and plan experiments to further elucidate these phenomena. For example, the effect of blood externally applied when inducing parthenogenesis in unfertilized frog's eggs by pricking needs further investigation. It is well known (Bataillon, 1912) that the presence of blood materially increases the percentage of haploid embryos reaching an advanced developmental stage. Tyler (1931) has described radially symmetrical parthenogenetic embryos of *Urechis*. He suggests that these received a diffuse activating stimulus, rather than a stimulus from one side (such as that provided in normal fertilization), and leading to formation of a bilaterally symmetrical embryo. Tyler (1941) suggests that parthenogenetic activating agents would best be applied in the form of a gradient. In line with these suggestions, we might postulate that the presence of the blood modifies the diffusion gradients set up through the point of puncture of the pricked frog's eggs, or modulates the stimulus to produce more normal embryos. One possible experimental test of this hypothesis is the study of the action, under similar circumstances, of various large molecules with isoelectric points near those of the several blood components.

Utilization of the Teorell diffusion effect as an explanation of the mechanism of localization of certain substances within the ovum does not necessarily imply that the diffusion agent enters the cell from outside. A diffusion gradient of substances leaving the cell, or diffusing from one region to another within the protoplasm, could produce a similar effect. The original polarization of the ovum, laid down during the early growth stages in the ovary, presumably sets up a polar difference in metabolically significant substances. This is one of the primary tenets of the axial gradient theory of C. M. Child. The products directly or indirectly resulting from these differences in metabolic activity may be free to diffuse from regions of higher to those of lower concentration. In so diffusing, under certain conditions, a Teorell diffusion effect may be established, thus secondarily inducing a movement of other charged particles. Gene products, diffusing from the nucleus or chromosomes during certain periods of cell activity, may similarly bring about movements of other substances.

A SPECIAL CASE: *Styela partita.* While the scheme outlined above (in terms of a polar defect) might account for segregation of materials along the polar axis, it could not, without addition or modification, account for a more complex type of oöplasmic segregation, such as that found in the egg of *Styela* and described by Conklin (1905). Upon fertilization of this ascidian egg, there is a primary segregation of materials resulting from a downflow of the yellow and clear substances from the animal toward the vegetal pole (Conklin, 1905, Figures 1–6). This active migration is completed within ten minutes after the entrance of the spermatozoon. Then the sperm nucleus moves to one side in the lower hemisphere, inaugurating a secondary segregation of materials to form the posterior yellow crescent (Conklin, 1905, Figures 8, 9, 13, 14, 15). Opposite this crescent, at the future anterior region of the egg, the light gray crescent arises. As a result of segregation, the animal hemisphere is occupied by clear protoplasm and the remainder of the vegetal hemisphere by dark gray yolk. In line with the polar defect theory outlined above, it would be tempting to suggest that the sperm entrance point serves as a second point of entrance for a diffusing substance which sets up the secondary bilateral pattern of the embryo, but such an explanation is inadequate. As Conklin (1905) demonstrated, the sperm nucleus does not always take the shortest path to the equator, but appears to move in a certain meridian. This seems to indicate that the path of the spermatozoon is determined by the structure of the cytoplasm. We must, thus, assume a bilateral orientation predelineated in the "framework" of the ground substance. Therefore, Harrison's (1945) diagram of the pattern of the ovum must be modified to include bilaterality as well as polarity, in order to be applicable to the case of the unfertilized *Styela* egg. The situation obtaining in the unfertilized *Styela* egg may perhaps best be visualized in terms of a bilateral liquid crystal structure extending throughout the cell (FIGURE 14), serving as the "framework" of the protoplasm, and leading to a directed diffusion. If directed diffusion leads to the establishment of a particular pattern of oöplasmic segregation, perhaps we may dispense entirely with the polar defect hypothesis. It is probable that segregation takes place in both the polar and antipolar fragments of cut *Cerebratulus* eggs. The antipolar fragment lacks the polar defect. Possibly, a liquid crystal structure with polar and bilateral orientation provides the requisite conditions.

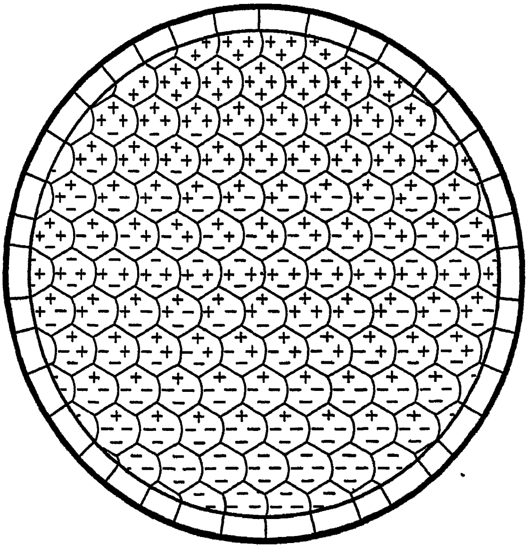


FIGURE 14. Diagram of hypothetical structure of an ovum, modifying HARRISON'S (1945) text figure 12 to include a cortex and a basic pattern of bilaterality in the lattice of the ground-substance. The lattice also provides a structural arrangement for the polarity of direction, while the distribution of the two kinds of particles (designated by + and −) within the interstices of the lattice indicates the polarity of stratification.

Undoubtedly, different patterns exist in the unfertilized eggs of different species of animals. In the frog's egg, it is probable that bilaterality is not determined until the time of fertilization. In the eggs of ascidians, it is apparently predetermined in the unfertilized egg. In the eggs of mollusks and annelids showing spiral cleavage, bilaterality may be determined at the time of fertilization (Just, 1912; Morgan and Tyler, 1930), but there is presumably an asymmetry of structure leading to spiral cleavage of the egg or its fragments, which is determined at an early precleavage stage.

The Significance of Oöplasmic Segregation. No more striking example of the significance of oöplasmic segregation can be found than in the case of differentiation without cleavage in the egg of *Chaetopterus*. After a short exposure to certain solutions (mixtures of 2½ molar potassium chloride and sea-water) both fertilized and unfertilized eggs pass through certain well-defined phases of segregation, as described by Lillie (1902, 1906) and by Brachet (1937). The yolk accumulates as a dense mass in the interior, and other granules or vacuoles assume a polar or peripheral position. FIGURE 15* shows a photomicrograph of one of these embryos, just before the differentiation of cilia. There is a remarkably clear-cut segregation of parts. Subsequently the peripheral protoplasm becomes ciliated and more vacuolated, so that the embryo resembles, in part at least, a trochophore with inner yolky endoderm, ciliated surface and equatorial band of vacuoles. Lillie's (1902) Figure 8 shows a normal trochophore and his Figures 1–7, etc., show the ciliated structures which develop without cleavage from unfertilized eggs. Embryos differentiating without cleavage do not develop an apical tuft. Lillie correlates this with the fact that the defect in the ectoplasm, where the endoplasm comes to the surface, is obliterated by the artificially induced flowing movements.

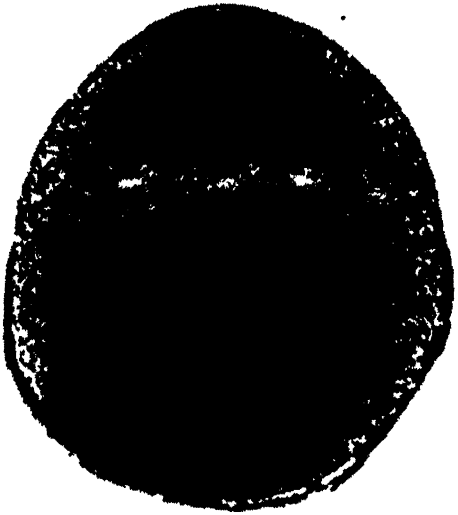


FIGURE 15. Photomicrograph of an unfertilized egg of *Chaetopterus pergamentaceus*, treated with a mixture of 2.5 M potassium chloride (10 parts) and sea-water (90 parts) for 64 minutes. Photographed 3 hours and 45 minutes after treatment. Note marked segregation of cytoplasmic components. A few hours later, cilia were differentiated at the surface. Magnification 660x.

In the same cultures,* some eggs are found in which the streaming movements of the interior protoplasm continue after the initial segregation has been accomplished. These amoeboid masses never differentiate the structures so characteristic of the other pseudo-larvae. Therefore, it appears that the proper degree of segregation must be maintained for differentiation to occur.

---

* See footnote on page 670.

It is my belief that oöplasmic segregation is of much greater significance in interpreting the development of egg-fragments than has been suspected previously. Studies of the development of egg-fragments, such as the classical experiments of E. B. Wilson (1904) on the egg of *Dentalium*, clearly indicate an association of specific embryonic potency with certain protoplasmic areas. For example, the lower polar area of the egg of *Dentalium* contains the material of the antipolar lobe. This region is definitely related to the formation of apical tuft and post-trochal regions. In *Dentalium*, there is, thus, a very early prelocalization of embryonic potency, associated with the very early visible segregation of the lower polar area.

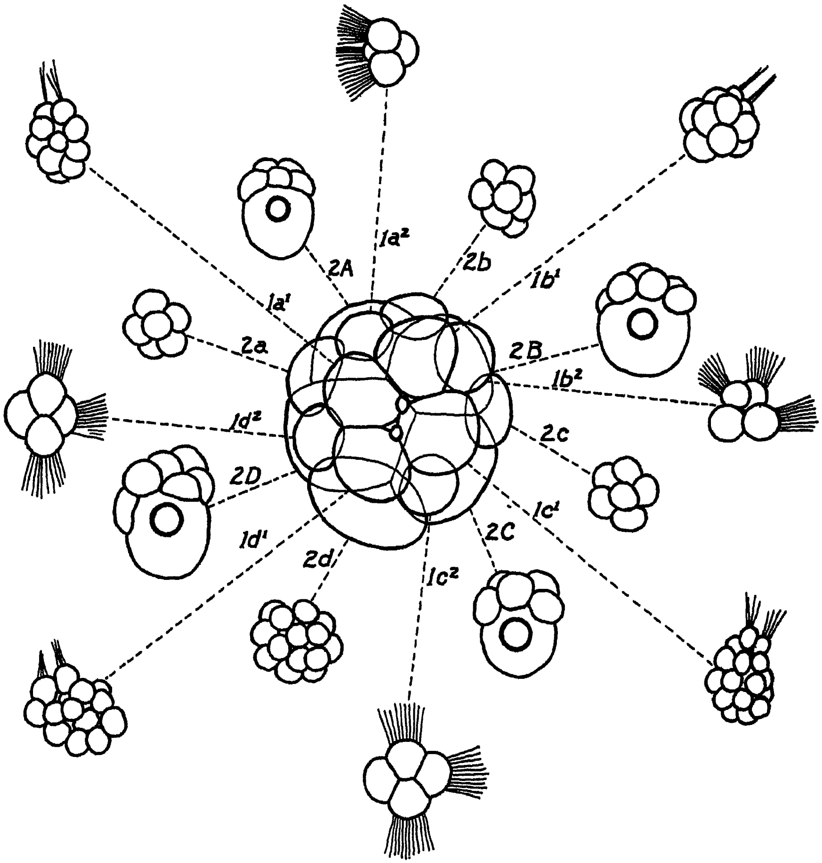


Figure 16. Differentiation of the blastomeres of the 16-cell stage of the *Nereis* egg after separation by dissection. (After Costello, 1945b.)

In the egg of *Nereis limbata*, it has been shown (Costello, 1945b) that, from the time of the first cleavage, the isolated blastomeres develop as partial embryos. This is demonstrated very clearly by isolating the blastomeres of the 16-cell stage (FIGURE 16), when only isolated trochoblasts ($1a^2$–$1d^2$) differentiate prototrochal cilia, only macromeres tend to gastrulate, etc. It is, therefore, of interest to inquire whether this cleavage mosaic is foreshadowed by a prelocalization extending back into the unsegmented egg. Cutting the *Nereis* egg across the equator, shortly before the first cleavage, produces polar fragments (upper fragments) containing the cleavage spindle, and antipolar fragments lacking spindle and nuclear materials. Only the polar fragments cleave (TABLE 1) and may produce essentially complete embryos (Costello, 1940a). Isolated

TABLE 1

DEVELOPMENT OF FRAGMENTS OF FERTILIZED *NEREIS* EGGS OBTAINED BY HORIZONTAL SECTION

	Number							
	Isolated	*Cleaved*	*Ciliated*	*Gastrulae*	*Proto. pigm.*	*Eyespots* (*1*)	(*2*)	*Anal pigm.*
Nucleated (polar)	25	20 (+3 abn.)	16	16	4 (+1?)	1 (+1?)		1
Non-nucleated (antipolar)	25	0	0	0	0	0		0
Whole denuded eggs	35	35	35	+	12	6	6	6

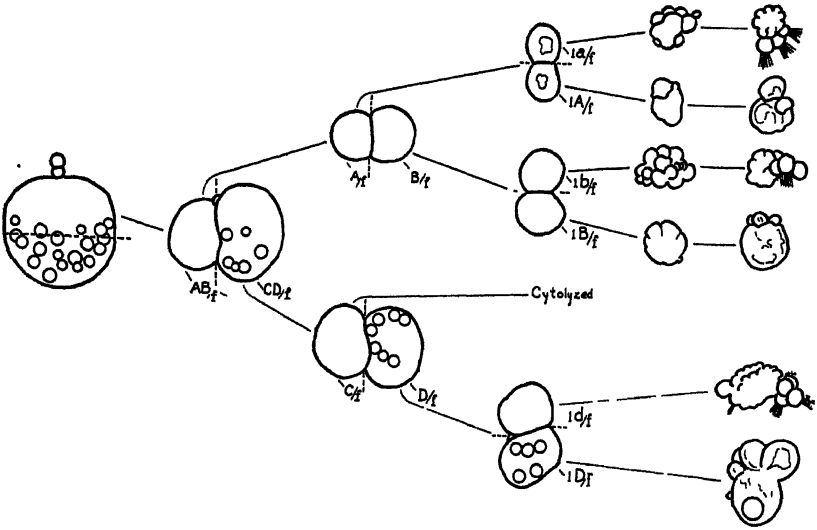


FIGURE 17. *Camera lucida* sketches of progressive blastomere separation of the cleavage products of an egg fragment of *Nereis*. The row of figures at the extreme right indicates the final products of differentiation of the surviving isolates. Compare with progressive blastomere separation of whole egg (COSTELLO, 1945b, Figure 4).

blastomeres of these egg-fragments develop in essentially the same manner as isolated blastomeres of whole eggs (FIGURE 17), clearly indicating that the prototrochal material cannot be at the extreme lower pole of the fragment, but occupies a position corresponding to its position in the whole egg (Costello, 1940b). Originally, I suggested three alternative hypotheses (FIGURE 18) to account for this result. First (FIGURE 18, a), it is possible that there is no prelocalization of the materials for prototroch formation (/) and for gastrulation (\) just prior to the first cleavage, when the egg-fragments were obtained. Secondly (FIGURE 18, b), it is possible that these materials are already segregated, but are segregated in the animal hemisphere, the lower half of the egg being essentially unimportant at this time. Thirdly (FIGURE 18, c), it is possible that the segregation of potencies has already occurred, but that the animal fragment is capable of regulating and produces the missing potencies out of other materials than those originally destined for these parts. I am now inclined to view the first hypothesis as the most likely. Since visible oöplasmic segregation in *Nereis* does not begin until after the formation of the polar bodies and is not well-advanced until just before the third cleavage, it is reasonable to suppose that germinal prelocalization is occurring simultaneously with visible segregation of the formed cytoplasmic inclusions.

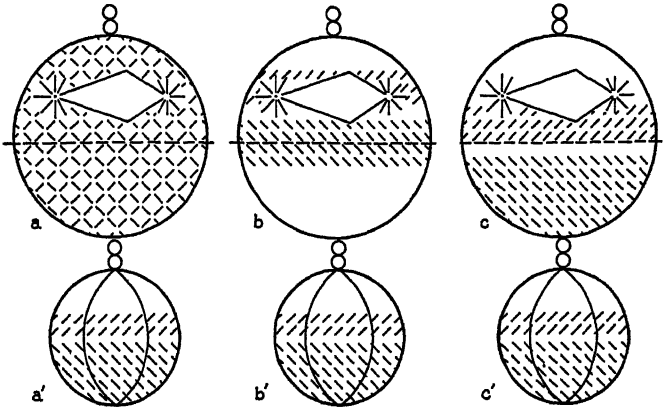


FIGURE 18. Diagrams of three alternative hypotheses concerning distribution of materials in whole eggs and egg-fragments (see text).

If a series of different egg species were studied, I believe that it would be possible to establish a correlation between the time of oöplasmic segregation and the degree of embryonic determination. However, we must bear in mind that centrifuging experiments on many forms have clearly indicated that there is no causal relation between visible particles, displaceable with centrifugal force, and morphogenetic values. For example, some of Spek's beautiful figures of *Nereis* eggs stained with neutral red and nile blue sulfate show a striking "bipolar differentiation." My own studies have shown that, when these stained *Nereis* eggs are subjected to an appropriate centrifugal force (FIGURE 19), all stained

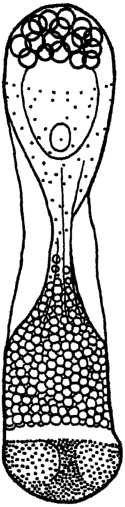


FIGURE 19. *Nereis* egg centrifuged 61 minutes at 66,000 times gravity. Sketched immediately, before change of shape. The strata, in order, proceeding from the centripetal pole, are: oil layer; hyaline zone surrounding upper portion of germinal vesicle; indistinct layer of fine granules, which stain with neutral red; broad stratum of yolk spheres; a second, narrow hyaline zone; centrifugal zone of heavy granules, including vortex of granules which stain deeply with neutral red.

materials are sedimented into definite strata. Such eggs may nevertheless develop normally. The "acid" and "alkaline" regions are not regions of acid or alkaline ground substance (hyaline protoplasm) but are regions in which the contained granules or vacuoles stain differentially. So far as we can ascertain by observation of stained and unstained eggs, there are no visible differences in the *hyaline protoplasm* of the different regions.

The conclusion that the visible cytoplasmic components of invertebrate eggs have no morphogenetic value has recently been questioned by Raven (1938) on the basis of experiments on the eggs of *Nereis* and *Chaetopterus*. Raven and Bretschneider (1942), using low centrifugal forces on the eggs of *Limnaea*, have objected to the conclusions of Conklin (1910) and Clement (1938). However, their objections appear to me to be without adequate foundation, since they ignore the fact that Clement (1938) obtained normal development of *hyaline fragments* of centrifuged *Physa* eggs. Harvey (1946) recently obtained plutei from the clear quarter of the *Arbacia* egg.

The apparent paradox between the results on the development of egg-fragments and of centrifuged eggs admits of easy solution. We need only assume that the mechanism of normal segregation segregates both *visible* formed inclusions and *invisible* morphogenetic substances. The invisible morphogenetic substances, such as hormones, enzymes, or the like, become associated with the hyaline protoplasmic base in the interstices between the granules or vacuoles. Subsequent centrifuging might easily displace the large visible particles without displacing the invisible substances which are associated with the protein framework of the cell. This conclusion is basically the same as that of Conklin (1931) resulting from centrifugation of the *Styela* egg. It is not the visible granules that have morphogenetic value, but the special, localized hyaline protoplasm with which these granules are normally associated.

The Teorell diffusion effect theoretically provides us with a mechanism which might bring about both a visible stratification of the suspended oöplasmic substances and a parallel but invisible segregation of the "formative stuffs." The primary requirement is that both types of materials be charged.

It has been suggested by Weiss (1939) that diffusible substances (such as inductors or hormones) can act only to bring out differences already existing in an embryo. He uses as an analogy the photographic developer which does not create the picture, but merely converts a latent image into a visible one. Weiss states (pp. 441–442), "It is inconceivable that a chemical agent diffusing indiscriminately through a body whose parts are all alike should ever be able to produce local differences. . . . the problem of organization can expect no elucidation from the study of the 'dark-room' phase of the process." I should like to point out here that a diffusible substance which possesses a difference of mobility of its dissociable parts could act, *via* the Teorell effect, to bring about a *segregation* of other charged particles and thereby actually create a pattern where none previously existed.* Two diffusion gradients of different types, originating from points 90 degrees apart, might give a more complex pattern. An interaction between the components of this pattern, once a certain threshold of concentration has been reached, might produce still more complex patterns.

Another example of a pattern which can be produced by diffusion of a substance into a gel lacking a preformed structural basis for the pattern is the well-known Liesegang phenomenon of colloid chemistry. When two solutions, capable of forming a relatively insoluble precipitate, are allowed to interact inside a gel, the precipitate, under certain conditions, is deposited in layers which are repeated in regular fashion. These are the Liesegang rings, or Liesegang stratifications. The explanation of this phenomenon, which involves diffusion in relation to the degree of local supersaturation, can be found in any good textbook of colloid chemistry.

---

* Obviously, diffusion from one side, or directed diffusion, is not *indiscriminate* diffusion. But one could visualize a pattern of segregation arising if an egg rests with one surface against the substratum, inhibiting diffusion from this side, and permitting free diffusion into the free surface from the external medium (*vide* work on the *Fucus* egg, summarized by Whitaker, 1940). The diffusion of the amphibian organizer from the chorda-mesoderm into the overlying ectoderm is no more indiscriminate. Organizer action, which appears to depend upon the competence of the reacting tissue, is probably a typical example of "developer action" similar to that described by Weiss for hormones.

In view of the lack of data on the physico-chemical conditions within the developing ovum, the diffusion effect theory of oöplasmic segregation is, at present, only a working hypothesis. I have suggested this hypothesis, not in the belief that it explains all the facts, but rather in the hope that a more adequate explanation of oöplasmic segregation may be evolved. It is hoped, furthermore, that this presentation may rekindle interest in the classical materials of experimental embryology, the eggs of the marine invertebrates.

Bibliography

Bataillon, E. 1912. La parthénogenèse des Amphibiens et la "fécondation chimique" de Loeb (Étude analytique). Ann. Sci. Nat. Zool. 9e Ser. 16: 249–307.

Brachet, J. 1937. La différenciation sans clivage dans l'oeuf de Chétoptère envisagée aux points de vue Cytologique et Métabolique. Arch. de Biol. 48: 561–589.

Child, C. M. 1940. Lithium and echinoderm exogastrulation: with a review of the physiological-gradient concept. Physiol. Zool. 13: 4–42.

Clement, A. C. 1938. The structure and development of centrifuged eggs and egg fragments of *Physa heterostropha.* J. Exp. Zool. 79: 435–460.

Conklin, E. G. 1905. The organization and cell-lineage of the ascidian egg. J. Acad. Nat. Sci. Phila. Ser. 2. 13: 1–119.

1910. The effects of centrifugal force upon the organization and development of the eggs of fresh water pulmonates. J. Exp. Zool. 9: 417–455.

1931. The development of centrifuged eggs of ascidians. J. Exp. Zool. 60: 1–119.

Costello, D. P. 1940a. Development of fragments of *Nereis* eggs. (Abstract) Anat. Rec. 78: (Suppl.) 133.

1940b. The development of isolated blastomeres of *Nereis* egg fragments. (Abstract) Anat. Rec. 78: (Suppl.) 133.

1945a. Segregation of oöplasmic constituents. J. Elisha Mitchell Sci. Soc. 61: 277–289.

1945b. Experimental studies of germinal localization in *Nereis*. I. The development of isolated blastomeres. J. Exp. Zool. 100: 19–66.

Harrison, R. G. 1945. Relations of symmetry in the developing embryo. Trans. Conn. Acad. Arts & Sci. 36: 277–330.

Harvey, E. B. 1946. Structure and development of the clear quarter of the *Arbacia punctulata* egg. J. Exp. Zool. 102: 253–275.

Herbst, C. 1892. Experimentelle Untersuchungen über den Einfluss der veränderten chemischen Zusammensetzung des umgebenden Mediums auf die Entwicklung der Tiere. I. Versuche an Seeigeleiern. Z. wiss. Zool. 55: 446–518.

Just, E. E. 1912. The relation of the first cleavage plane to the entrance point of the sperm. Biol. Bull. 22: 239–252.

Lillie, F. R. 1902. Differentiation without cleavage in the egg of the annelid *Chaetopterus pergamentaceus.* Arch. Entw. 14: 477–499.

1906. Observations and experiments concerning the elementary phenomena of embryonic development in *Chaetopterus.* J. Exp. Zool. 3: 153–268.

Mead, A. D. 1898. The rate of cell-division and the function of the centrosome. Biological Lectures of the M. B. L., Woods Hole: 203–218.

Morgan, T. H., & A. Tyler. 1930. The point of entrance of the spermatozoön in relation to the orientation of the embryo in eggs with spiral cleavage. Biol. Bull. 58: 59–73.

Pease, D. C. 1942. Echinoderm bilateral determination in chemical concentration gradients. II. The effects of azide, pilocarpine, pyocyanine, diamine, cysteine, glutathione and lithium. J. Exp. Zool. 89: 329–345.

Raven, C. P. 1938. Experimentelle Untersuchungen über die "bipolare Differenzierung" des Polychaeten- und Molluskeneies. Acta Neerlandica Morphol. 1: 337–357.

RAVEN, C. P., & L. H. BRETSCHNEIDER. 1942. The effect of centrifugal force upon the eggs of *Limnaea stagnalis* L. Arch. Néerland. Zool. 6: 255–278.

SPEK, J. 1930. Zustandsänderungen der Plasmakolloide bei Befruchtung und Entwicklung des *Nereis*-Eies. Protoplasma 9: 370–427.

1934a. Über die bipolare Differenzierung der Eizellen von *Nereis limbata* und *Chaetopterus pergamentaceus*. Protoplasma 21: 394–405.

1934b. Die Reaktion der Protoplasmakomponenten des *Asterias*-Eies. Protoplasma 21: 561–576.

1938. Studien über die Polarität der Larven der Kalkschwämme. Protoplasma 30: 352–372.

TEORELL, T. 1935a. Studies of the "diffusion effect" upon ionic distribution. I. Some theoretical considerations. Proc. Nat. Acad. Sci. 21: 152–161.

1935b. Some aspects of elektrolyte diffusion. (Abstract) Biol. Bull. 69: 331.

1935c. On an arrangement for studying the conditions within diffusion layers. Science 81: 491.

1937. Studies of the diffusion effect upon ionic distribution. II. Experiments on ionic accumulation. J. Gen. Physiol. 21: 107–122.

TYLER, A. 1931. The production of normal embryos by artificial parthenogenesis in the echiuroid, *Urechis*. Biol. Bull. 60: 187–211.

1941. Artificial parthenogenesis. Biol. Rev. 16: 291–336.

WEISS, P. 1939. Principles of Development. Henry Holt & Co. New York.

WHITAKER, D. M. 1940. Physical factors of growth. Growth (Suppl.) 75–88.

WILSON, E. B. 1892. The cell-lineage of *Nereis*. J. Morphol. 6: 361–480.

1904. Experimental studies on germinal localization. I. The germ-regions in the egg of *Dentalium*. J. Exp. Zool. 1: 1–72.

Discussion of the Paper

DR. ROBERT CHAMBERS (*New York University, New York, N. Y.*):

I was interested in Dr. Costello's introducing a diffusion gradient model as a hypothesis to explain orientation in an egg. I only wish he could have taken time to present more, and in greater detail, from the great wealth of his own experimental observations.

Concerning the existence of a polar differentiation which persists from the ovarian egg, I wish to mention the egg of the *Cerebratulus*. The excentrically placed germinal vesicle maintains the same relative position as that of the nucleus of the cells of the germinal epithelium, while the micropyle, which is an interruption of the vitelline membrane, as it expands on contact with sea-water, lies at the opposite pole where the epithelial cell had separated from its substrate. Hence, under normal conditions and in the absence of external disturbing factors, the polar bodies arise almost exactly opposite the spot where the sperm is most likely to enter. Insemination can occur anywhere over the surface of the egg, but the growing aster, once the sperm has entered, brings the sperm-head into the position where nature had originally intended it to be. Can it be that the gelating monaster, with the streaming entailed in its growth, has something to do with a reorientation of previously dislocated basic patterns? We know that the development of totipotent egg-fragments is also preceded by the formation of a monaster. May not this phenomenon produce the required polar orientation in each fragment?

Dr. D. P. Costello:

Theoretically, it seems possible, as Dr. Chambers suggests, that the streaming movements associated with the growth of the sperm aster may orient substances within the egg, to give a polarity of stratification, though not a polarity of direction. The latter is probably laid down in the egg during its earliest oögonial history. However, if we postulate this activity as a characteristic of sperm asters, we encounter the difficulty of explaining orientations which bear no constant relation to the position of the sperm path, as has been pointed out for the case of the secondary bilateral segregation in the egg of *Styela.* Rashevsky has suggested the possibility that the centrioles of asters are diffusion centers, bringing about an orientation of chromosomes in the mitotic figure, but there is no direct evidence for this. He has not discussed centrioles in relation to oöplasmic segregation.

Dr. A. M. Shanes (*New York University, College of Dentistry, New York, N. Y.**):

Dr. Costello has suggested that Teorell's theoretical approach to the "diffusion effect," devised to account for ionic gradients, is applicable to the distribution and movement of particles in the fertilized egg. It may be desirable to call attention to a limitation of the theory as developed by Teorell and its possible significance.

Teorell assumes that diffusion potentials set up by the continuous diffusion of ionized substances are responsible, in the steady state, for the differential distribution of ions on either side of any boundary which serves to support such diffusion gradients. Unfortunately, under conditions of rather high electrolyte concentration such as exist within most if not all cells, such diffusion potentials would be insignificant. Thus, Dr. Osterhout set up a model in which CO_2 served as a source of diffusing ions across a non-aqueous layer. A differential distribution of ions occurred, but the expected potentials arose as ionic changes proceeded rather than before. The potentials were obviously the result rather than the cause of the ionic transfer, and this even though hydrogen ions (the most mobile and therefore the best from the standpoint of large diffusion potentials) were involved and their gradient very high.

Dr. Osterhout's results are easily interpreted from the standpoint of the Donnan equilibrium, for bicarbonate ions were retained within the "cell" while hydrogen ions exchanged with other "extracellular" small cations. Thus bicarbonate constituted the indiffusible ions required for a Donnan equilibrium; the ionic gradients and associated potentials would thus develop as actually observed.

It is difficult to see, in view of such results, how a diffusion effect could cause the movement of microscopically visible particles, particularly with the speed and over distances as great as those which have been described for some eggs. Even if molecules as small as proteins were concerned, the porosity of the protoplasm necessary for their migration would eliminate the possibility of ion retention such as seems to be required for the diffusion effect.

---

* Present address: Georgetown University, School of Medicine, Washington, D. C.

THE ORGANIZATION OF THE AMPHIBIAN EGG DURING FERTILIZATION AND CLEAVAGE

By G. FANKHAUSER
Princeton University, Princeton, N. J.

Introduction

In contrast to the beautifully transparent or semi-transparent eggs of invertebrates and ascidians reviewed by Costello, where processes of oöplasmic segregation may be followed under the microscope, the eggs of amphibians are perfectly opaque. Following fertilization, one can detect sufficient changes at the egg surface to be sure that more important rearrangements of materials must be taking place inside. However, when the egg is fixed and sectioned, one merely sees a heavy suspension of yolk platelets, with a small amount of cytoplasm scattered among these presumably inert materials. There is little to attract our immediate attention and to arouse our curiosity.

It is no wonder that the study of the organization of the amphibian egg in fertilization and cleavage was long neglected. Well in the foreground during the early days of *Entwicklungsmechanik*, in the hands of Born, Schultze, Roux, and Spemann, it was pushed into the background by the sensational success of the transplantation method that may be applied to the egg in slightly more advanced stages. During the past few years, there has been a revival of interest in the initial stages of development. Obviously, to trace the origin of the already complex organization at the blastula and gastrula stages is as important a task today as it ever was (*cf.* Harrison, 1945).

The unsatisfactory state of our knowledge of the egg in its earliest stages is vividly portrayed by the fact that the revival of interest at once led to an animated controversy between different workers. Dalcq and Pasteels assume a relatively simple organization of the ovum at fertilization which becomes gradually more complex during cleavage. Lehmann, on the other hand, is convinced that the condition of the unsegmented egg is very similar to that of the beginning gastrula, with little change during segmentation. Although the two viewpoints appear to be fundamentally different, they are likely to converge as more facts become known.

To make matters worse, the professional language of the embryologist is full of ambiguous and, perhaps, outmoded terms that are bound to foster misunderstandings. Several eminent investigators have warned us against the dangers inherent in the use of such expressions as "determination," "segregation," "organ-forming substances," "organizers," etc.

The list of terms "the use of which is not recommended," to borrow Needham's phrase (1942), is growing all the time and may soon include most of those that have become dear to us through long association and, therefore, seem to be full of meaning, although it now appears that this meaning may not be exactly the same for all workers in the field. It seems doubtful that the solution to these problems of vocabulary will be found in the introduction of a whole series of new terms. While some, if generally accepted, may help to clarify, others are bound to add to the confusion. As far as the organization of the ovum is concerned, the primary need is for more and better established facts of sufficient convincing power to bring about more general agreement among different workers on the factual level.

Under these circumstances, it may be most profitable to review the overall picture of the organization of the amphibian ovum as it stands today, to point out the few accomplishments and the great gaps in our knowledge that still exist. Such a review should comprise (*1*) the observations on the visible organization of the egg, including both nuclear and cytoplasmic phenomena; (*2*) the experimental tests of the invisible organization of the cytoplasm; and (*3*) the results of experimental analysis of the nucleus in so far as they concern our problem.

Visible Organization of the Amphibian Egg

Nuclear Phenomena. At the very beginning of its career, the egg of almost all salamanders passes through a crisis which might well prove to be fatal if it were not for the existence of a special compensating mechanism. In frogs and toads, fertilization is normally monospermic. If, under laboratory conditions, two or more spermatozoa are allowed to enter the egg, the supernumerary sperm nuclei divide independently, at the same time as the diploid fusion nucleus. Cleavage is abnormal, leading to the formation of haploid, diploid, and mixed cells, and development comes to an end in embryonic or early larval stages (Brachet, 1910, 1912; Herlant, 1911).

In the majority of urodeles, on the other hand, fertilization is normally polyspermic. Up to ten or more spermatozoa may enter the egg, depending on the species. During the first three hours, the internal developments are similar to those in polyspermic frog's eggs and seem to predict certain disaster. However, at the critical time, as the principal sperm nucleus unites with the egg nucleus, the accessory sperm nuclei begin to show signs of degeneration. They may go on to prophase and even release the chromosomes, but the sperm asters do not divide, and the remnants of the accessory nuclei are soon pushed out of the way by the large asters of the diploid mitotic system (FIGURE 1; Fankhauser 1932a; Fankhauser and Moore, 1941a). The nature of the inhibiting factor is not known. It seems to spread from the vicinity of the dominant nuclear system. In any case, there must be a change in the cytoplasm

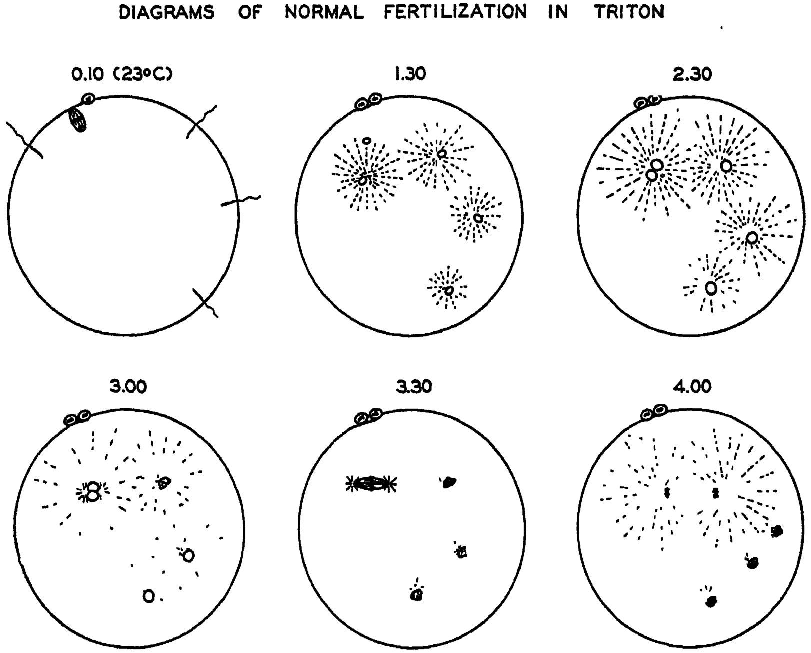


FIGURE 1. Diagrams of polyspermy in the European Newt, *Triton*. The egg is represented in side view, animal pole up; the nuclei, maturation spindle, and cleavage spindle at metaphase are greatly enlarged relative to the size of the whole egg. The asters are not enlarged.

10 minutes after insemination (*at 23° C.*): metaphase of second maturation division, penetration of four spermatozoa.

1 hour 30 minutes: second polar body given off; small egg nucleus moves toward nearest sperm nucleus which will become the principal sperm nucleus. All accessory sperm nuclei develop normally. Differences in the size of the sperm asters are related to the amount of active cytoplasm present in the animal and vegetal hemispheres.

2 hours 30 minutes: egg nucleus and principal sperm nucleus in contact. Maximum development of sperm asters. Accessory sperm nuclei still normal.

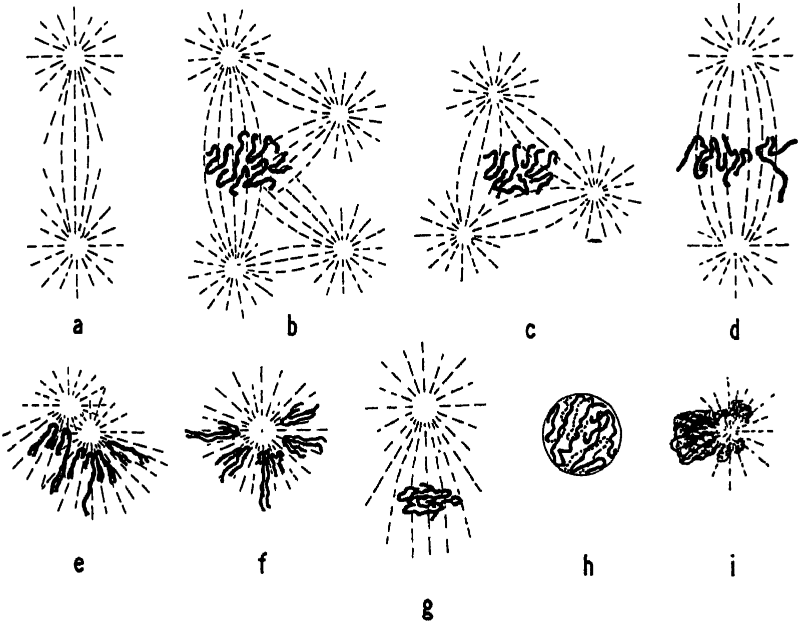
3 hours: fusion of egg and principal sperm nucleus; fading of sperm asters. Two small asters appear in center of principal sperm aster. Accessory sperm asters remain undivided. Accessory sperm nucleus nearest fusion nucleus shows signs of degeneration.

3 hours 30 minutes: metaphase of first cleavage mitosis. All accessory sperm nuclei degenerating. Note eccentric position of cleavage spindle.

4 hours: early telophase of first cleavage mitosis. Note growth of asters at poles of first cleavage spindle which tends to center the mitotic figure and to push the remnants of accessory sperm nuclei out of the animal hemisphere.

surrounding the accessory sperm complexes which blocks their division.

We easily forget that both nucleus and cytoplasm are also involved in the fulfillment of another prerequisite of normal development, *viz.*, the establishment of a mitotic apparatus that will assure normal cleavage of the egg. Under normal conditions, the cycles of the nuclear and cytoplasmic components of the mitotic system are perfectly coordinated. However, if we prevent the union of the egg nucleus and the principal sperm nucleus, by dividing the fertilized egg into two parts, or by eliminating one of the nuclei, the chromosomal and centrosomal cycles are frequently thrown out of gear. In fragments of salamander eggs, and in whole eggs following the removal of the egg chromosomes, the isolated sperm nuclei show a whole series of abnormal mitotic phenomena: absence or inactivity of the division center leading to a monaster; delayed division of the center; precocious division producing multipolar figures; and, finally, the cytoplasmic component may go through the paces all alone and still be able to induce division of the cell body (FIGURE 2; Fankhauser, 1934a; Fankhauser and Moore, 1941b).



BEHAVIOR OF SPERM NUCLEUS IN ANDROMEROGONY IN *TRITON*

FIGURE 2. Diagrams of various types of mitosis found in androgenetic egg fragments of *Triton*. In the absence of the egg nucleus, accessory sperm nuclei may divide. Frequently, the chromosomal and centrosomal cycles are out of gear. Several types of mitosis may be found in a single egg fragment.

a, division of aster in complete absence of chromosomes.

b and c, precocious division of sperm aster producing tetra- and tripolar figures.

d, normal mitosis.

e, delayed division of aster, at a time when the chromosomes had oriented their attachment points toward the single aster. Whole (split) metaphase chromosomes are distributed irregularly between the two poles.

f, typical monocentric mitosis (monaster); aster remains undivided, chromosomes divide normally.

g, abortive monaster (chromosomes fail to orient their centromeres toward the single pole and do not divide).

h, sperm nucleus in prophase, no aster (later stages of "anastral" mitosis of sperm nuclei have not been found).

i, degeneration of sperm nucleus without attempt at division.

The egg nucleus, when isolated in a fragment of the fertilized egg, is in a more difficult position since it is not normally associated with an active division center. A centrosome may arise, but it usually fails to divide. If no centrosome appears, the egg nucleus may remain in late prophase and grow to extraordinary size, or it may divide repeatedly by an anastral mitosis that looks as if it had been borrowed from a plant cell (FIGURE 3). It is important to note that this anastral form of mitosis also may induce cleavage (Fankhauser, 1937). Cell division itself, the most basic process of development, is an example of a phenomenon that was known, in the days of classical developmental mechanics, under the perhaps inadequate but illustrative name, "multiple assurance."

BEHAVIOR OF EGG NUCLEUS IN GYNOMEROGONY IN *TRITON*

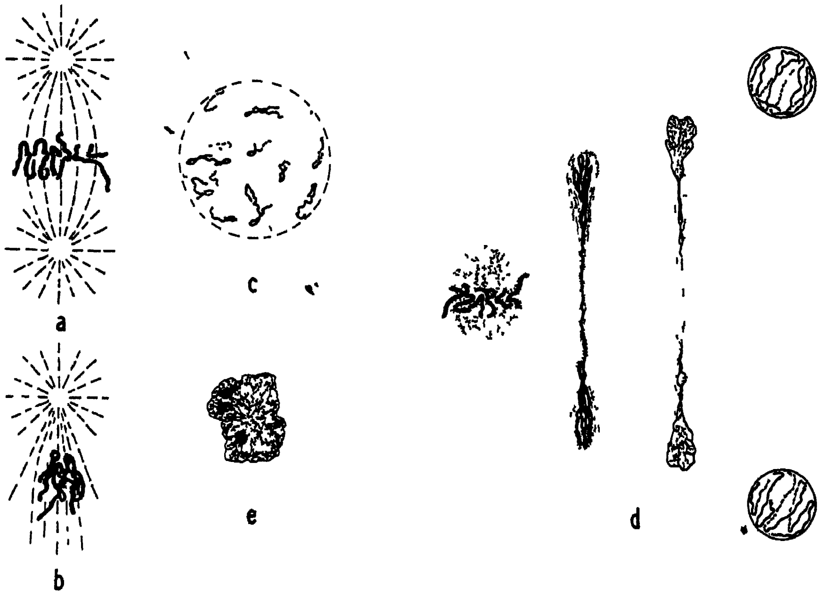


FIGURE 3. Diagrams of types of mitosis in gynogenetic egg fragments of *Triton* (fragments of fertilized eggs containing the egg nucleus alone).

a, normal mitosis of egg nucleus (very rare).

b, abortive monaster.

c, "permanent" prophase with continued swelling of nucleus (twelve chromosomes, the haploid number in *Triton*, may be counted).

d, four stages in anastral mitosis; *metaphase:* with short "spindle" formed by "half-spindle components" derived from nuclear sap; *anaphase:* stretching of spindle material, incomplete separation of chromosomes; *telophase:* disappearance of spindle material (note similarity to amitosis); *prophase* of following mitotic cycle.

e, degeneration of egg nucleus without attempt at division.

In this connection, another observation should be mentioned which clearly shows how particular conditions in the ooplasm may control the behavior of the nucleus. Following A. Brachet's (1922) discovery of the peculiar reactions of the spermatozoa in immature sea urchin eggs, Bataillon (1929, 1934) described similar phenomena in eggs of *Hyla* and *Triton*. In immature eggs which contain the first maturation spindle, the spermatozoa do not form vesicular nuclei associated with large asters. Precociously, small chromosomes are released which are located on a single aster or in a small truncated spindle simulating the maturation spindle. The mitotic figures remain blocked in this condition until the egg dies. It is interesting that the same reaction may be obtained in mature eggs if these are first made "immature" again by a two-hour exposure to carbon dioxide.

Cytoplasmic Phenomena. The primary polarity of the egg is clearly expressed in the ovarian eggs of most amphibians in the distribution of superficial pigment and of the yolk platelets. From what we know of the origin of the animal-vegetal axis in eggs of invertebrates, it seems likely that it is present in the very young oocyte but initiated, or at least reoriented, by some factor in the ovarian environment. It should be emphasized, however, that the existence of such an external factor has never been demonstrated satisfactorily, contrary to the statements made in some textbooks.

Dorso-ventral polarity, or bilateral symmetry, becomes visible in many amphibian eggs after fertilization, when the grey crescent, which marks the position of the future dorsal lip, appears on one side, at or below the equator. Ancel and Vintemberger (1933) have shown that, in *Rana fusca*, the formation of the crescent is connected with extensive shifts of materials at the egg surface.

In some species, even unfertilized eggs may show bilateral symmetry very clearly. According to Pasteels (1937), certain egg batches of *Rana esculenta* show an oblique pigment boundary, higher on the future dorsal side. The same phenomenon had been described earlier in the axolotl, by Banki (1927, 1929). During the formation of the grey crescent, the whole egg seems to rotate, raising the dorsal side still further. Ancel and Vintemberger (1933) claim that this "rotation" is a purely cortical phenomenon and does not involve the whole mass of the egg. Banki (1929) also applied vital stain marks to the axolotl egg immediately after fertilization. Those on the ventral side remained localized, those on the dorsal side spread in the direction of the median plane during the formation of the grey crescent.

If we are, thus, certain that bilateral symmetry may be present in the egg before fertilization, we also know that the median plane may still be shifted after insemination. In some species, like *Rana fusca*, the position of the point of entrance of the sperm may have an important orienting influence, so that the grey crescent in most cases forms on the opposite side of the egg, as shown most recently by Ancel and Vintemberger (1938d). In other anuran species, there is no constant relation between the point of sperm entry and the median plane (*Rana esculenta*, *Discoglossus*; Pasteels, 1937, 1938). In urodeles, polyspermy would prevent a simple relationship. Even in selected monospermic eggs, the planes of fertilization and of bilateral symmetry do not coincide (*Cryptobranchus*, Smith, 1922; axolotl, Vogt, 1926; Banki, 1927).

The final position of the median plane may also be changed by an artificially imposed rotation of the egg. Ancel and Vintemberger (1938a) placed unfertilized eggs of *Rana fusca* on a slide in an oblique position, with the vegetal pole raised 135°. After fertilization and expulsion of the perivitelline fluid, the eggs could respond to gravity and rotate until they reached the normal equilibrium position. When the grey crescent appeared, it was almost always located on the "descending side" of the egg, *i.e.*, on that side over which the vegetal hemisphere descended during equilibration of the egg. The same result was obtained when the rotation was imposed as late as 60 minutes after fertilization. Beyond this time, it was without effect (Ancel and Vintemberger, 1938b, c). Further experiments showed that the effect of the first rotation may be reversed by a second rotation in the opposite direction, and that the orienting influence of the spermatozoon may be canceled by subsequent rotation (1938e). The observations were extended to *Rana esculenta* and *Triton alpestris* (1938f) and, for *Rana esculenta*, confirmed by Pasteels (1938). There is no doubt that, during the first hour following fertilization, an artificially imposed rotation of the egg can reorient the plane of symmetry. Later on, as the grey crescent begins to form, this treatment is no longer effective. It seems as if an easily displaced substance had now become fixed in a certain area.

Microscopical study of sections through eggs before or after fertilization has long been neglected. Banki (1929) cut axolotl eggs in half along the median plane and saw a broad cortical zone in the animal hemisphere, surrounding a brownish central mass, while the vegetal hemisphere was largely occupied by white, heavy yolk. During formation of the grey crescent, the cortical zone spread out considerably on the dorsal side and formed a thin layer.

More recently, Lehmann (1941, 1945) cut formalin-fixed eggs of the axolotl in two and described a ring of pigmented, sub-cortical or "marginal" plasm in the animal hemisphere which he identifies with the marginal zone of the gastrula, *i.e.*, with the presumptive notochord-mesoderm area. This conclusion is based on the observation that the ring of marginal plasm may be slightly wider on one side. The investigation seems to be of a preliminary nature and should be extended to eggs of various species of amphibians before it can support far-reaching theoretical conclusions.

Undoubtedly, it would be much easier to understand the organization of the egg at fertilization if it could be traced to its origins during the growth period in the ovary which, in amphibians, may last for weeks or months. A highly promising beginning has been made with the aid of histochemical tests, particularly in the hands of J. Brachet (1944). The so-called plasmal reaction, first applied to eggs of axolotls by Voss, indicates the distribution of a special type of phosphatide which is frequently associated with ribose-containing nucleoproteins. In the very young oocyte, the reaction is limited to the "yolk nucleus," a concentration of mitochondria, which later breaks up and forms a peripheral ring. It is interesting that the loci of synthesis of both lipids and proteins coincide with the areas containing plasmalogen.

Ribose-nucleic acid is always present in the nucleoli of the germinal vesicle; in young oocytes, the whole cytoplasm gives an intense reaction; later on, it is largely limited to a perinuclear ring. The germinal vesicle of the full-grown oocyte shows a high concentration of sulfhydryl compounds in the nuclear sap. Following the breakdown of the nuclear membrane, these compounds occupy roughly the same area as the ribose-nucleic acid. These substances spread out from the animal pole, descending, at first, more rapidly on one side, which seems to be the future dorsal side. At the beginning of gastrulation, they are concentrated particularly at the animal pole and in the dorsal lip of the blastopore. It is very tempting to connect the peculiar distribution of ribose-nucleic acid compounds with the origin of bilateral symmetry, although Brachet himself strongly emphasizes the hypothetical character of such a connection.

Experimental Tests of Invisible Organization of Egg Cytoplasm

Maps of Organ-Forming Territories. Although we are well informed of the prospective significance of the various regions of the blastula and gastrula, corresponding maps of the unsegmented egg have not been published so far, perhaps because of technical difficulties involved in applying vital stain marks at sufficiently early stages, before the cortical shifts described by Ancel and Vintemberger begin. Such maps would be extremely valuable in the analysis of the rearrangements of materials that take place after fertilization and during early cleavage, as is clearly shown by Banki's observations and by the studies on ingression which are discussed in detail in the paper by Dr. Nicholas.

Effects of Gravity and Centrifuging. The effects of gravity on amphibian eggs have been studied repeatedly since the days of Pflüger, Schultze, and Born, who turned frog's eggs upside down and forced them to develop in this inverted position, either by compressing them between two slides, or by preventing the swelling of the jelly and the formation of the normal perivitelline space. The resulting embryos were abnormal, largely because of disturbances of gastrulation which often began at two different points. The observations have always been rather difficult to interpret because we do not know in sufficient detail how the various substances within the egg are redistributed under the action of gravity. Pasteels (1941), who recently inverted eggs at the time of fertilization, also seems to have limited his observations largely to living eggs. On that basis, he constructed optical sections which indicate that the variable results may be explained by varying degrees of descent of the heavy yolk. In most cases, two blastopores are formed; no normal embryos appear, probably because gravity alone is unable, in the time available, to assure a complete reversal of the original orientation.

However, if inverted eggs are centrifuged at 460 gravities for from one to five minutes, a complete reversal of the original animal-vegetal (= antero-posterior) axis may take place. Gastrulation is unitary, and about one-half of the resulting embryos are perfectly normal, although their cephalic end corresponds to the original vegetal pole of the egg. The complete reversal of the primary polarity of the egg by mild centrifuging makes it appear unlikely that polarity is determined primarily by properties of the egg surface, since the cortex does not seem to be greatly affected by inversion or centrifuging.

Pasteels (1940b) also centrifuged eggs of *Rana fusca* which were free to rotate and to orient themselves with the animal pole pointing centripetally. Following centrifugation of two to three minutes, at 460 gravities, gastrulation was often normal. In spite of this fact, the axial organs were sometimes completely absent in all or part of the embryo. In some cases, the tail still contained an axial mass of recognizable somites while the trunk was completely unorganized beyond the formation of a mantle of mesoderm. Other embryos were completely anaxial although they differentiated numerous blood cells.

If eggs are centrifuged after formation of the grey crescent, fewer embryos survive, but the resulting tadpoles may be completely normal. The formation of the axial organs, or organogenesis in general, is no longer inhibited. From these observations Pasteels concludes that, at the time of fertilization, the egg contains a substance of relatively low specific gravity which is a precursor of the active principle of the organization center, called organisin. Perhaps this precursor is identical with the ribose-nucleic acid compounds which Brachet found in higher concentration on the dorsal side of the egg. Centrifuged eggs actually show these compounds concentrated at the animal pole. Once the grey crescent has been formed, the precursor becomes fixed in a more cortical area and cannot be displaced by centrifuging.

Pasteels (1940a), repeating still another old experiment, rotated eggs, following the formation of the grey crescent, through 135° and forced them to develop in this position by compression between two slides. The rearrangement of the materials is more uniform than after complete inversion, since the heavy yolk tends to sink down along one side of the egg. However, in its new position, it will have various spatial relations to the old grey crescent area. While the blastopore always appears at the boundary of the heavy yolk mass, its position is also influenced by the original location of the grey crescent. Very often, the point at which invagination begins seems to represent a compromise between these two forces.

These experiments form the basis of a theory of development (Dalcq and Pasteels, 1937, 1938; Dalcq, 1938) which recognizes two important features of the egg at fertilization: an animal-vegetative vitelline gradient, involving the whole mass of the egg, and a dorsal field limited to the cortex. As cleavage proceeds, an interaction between these two components takes place which is visualized as a simple chemical reaction producing different results at different points, because of the varying concentration of the vitelline substance "V" and the cortical substance "C." The products of this reaction would then create a field of "morphogenetic potentials" which, in turn, would determine the fate of various regions of the embryo.

This theory, which in its details is somewhat more complex than the brief résumé indicates, has the great merit of focusing our attention on the yolk material which is either not as inert as was formerly believed, or always associated with some other, active material. Also, it seems highly probable that interactions between the vitelline gradient and a more superficial dorsal "field" play an important role in the earliest stages of development. The theory, which undoubtedly will be subjected to more experimental tests, has been criticized by Rotmann (1943) and Lehmann (1945). The latter questions the existence of a true yolk gradient in the amphibian egg and claims that distinct masses of heavier and lighter yolk are present. In his opinion, the dorsal field is not cortical, but sub-cortical or "marginal." It is very likely that continued investigations will bring about a *rapprochement* of these conflicting views.

Isolation of Parts of Egg. The invisible organization of the cytoplasm of the egg before and during cleavage may also be studied by isolating parts of the egg, to test their developmental potencies. The interpretation of the results obtained by such methods is somewhat simpler since it may be formulated in terms of the organizer concept.

Spemann (1902) showed that eggs of *Triton* in early gastrula stages may be constricted within the jelly, by means of a loop of fine hair. When the loop divides the blastopore symmetrically, *i.e.*, when the plane of constriction coincides with the median plane of the embryo, complete twins are formed. When the gastrula is constricted in the frontal plane, so that the dorsal and ventral halves are isolated, the dorsal embryo alone develops normally; the ventral one may gastrulate but fails to form any axial organs (somites, notochord, and neural tube).

Similar results are obtained by dividing the blastula or earlier cleavage stages, or by isolating the two blastomeres at the two-cell stage (Spemann, 1901; Spemann and Falkenberg, 1919; Ruud and Spemann, 1922; Ruud, 1925). It is, thus, necessary to conclude that the dorsal and ventral sides of the egg are already different from one another at this early stage. Spemann was inclined to believe that the future center of organization becomes localized on the dorsal side of the egg before the first cleavage, perhaps at the same time as the grey crescent forms in eggs of other species of amphibians, *i.e.*, about one to two hours after fertilization.

However, if eggs of *Triton palmatus* are constricted as early as twenty minutes after fertilization, we still obtain the same results: about one-fifth of the fragments containing the diploid cleavage nucleus give rise to typical ventral embryos (Fankhauser, 1930b), a similar proportion as that produced in experiments on the two-cell stage. The ventral embryos may survive for several days and form mesenchyme, blood cells and pigment cells, but no trace of notochord, myotomes, or neural tissue ever appears (FIGURE 4). This demonstrates that an important difference

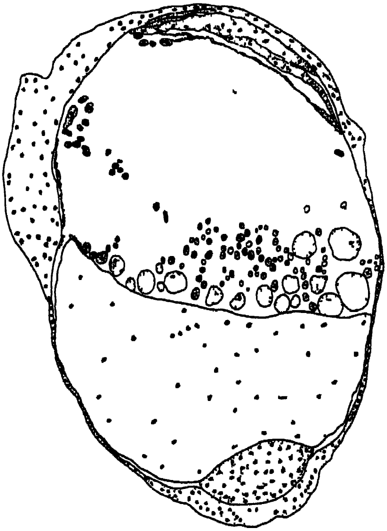


FIGURE 4 Section, along animal-vegetal axis, through typical ventral embryo, developed from a diploid egg fragment of *Triton palmatus* 6 days old (control embryo shows well differentiated neural tube, gill buds, and indication of forelimb buds) The egg was constricted twenty minutes after insemination Note presence of mesenchyme near animal pole, blood island (a typically "ventral" differentiation) near vegetal pole, irregular thickening of ectoderm, complete absence of axial organs

in developmental potencies exists between the dorsal and ventral sides of the egg within a few minutes from fertilization. The formation of the grey crescent seems to be a secondary phenomenon which makes this difference clearly visible in the eggs of some species. It is possible that the basic dorso-ventral differentiation is present before fertilization. However, attempts to obtain egg fragments by constriction of unfertilized eggs in the posterior part of the oviducts have not been successful so far.

There is no direct evidence that the dorsal area of the unsegmented egg already has all the properties of the center of organization of the blastula or gastrula, *e.g.*, that, on transplantation, it could at once induce formation of a secondary embryonic axis. Inductions have been obtained by bringing unsegmented eggs or ¼ blastomeres in contact with ectoderm of young gastrulae for 30 to 35 hours (Mayer, 1939). However, during this interval, the eggs developed at least to the blastula stage so that it is not possible to ascertain the exact time at which inductive powers arise. As far as the analysis of the cytoplasmic organization is concerned, the fact that the dorsal half of the egg has the capacity to form an organization center while the ventral half does not, is in itself important enough.

Experiments on the two-cell stage have the advantage that both cells develop with identical, diploid nuclei and may serve as mutual controls. If the two halves of the unsegmented egg are separated completely, both

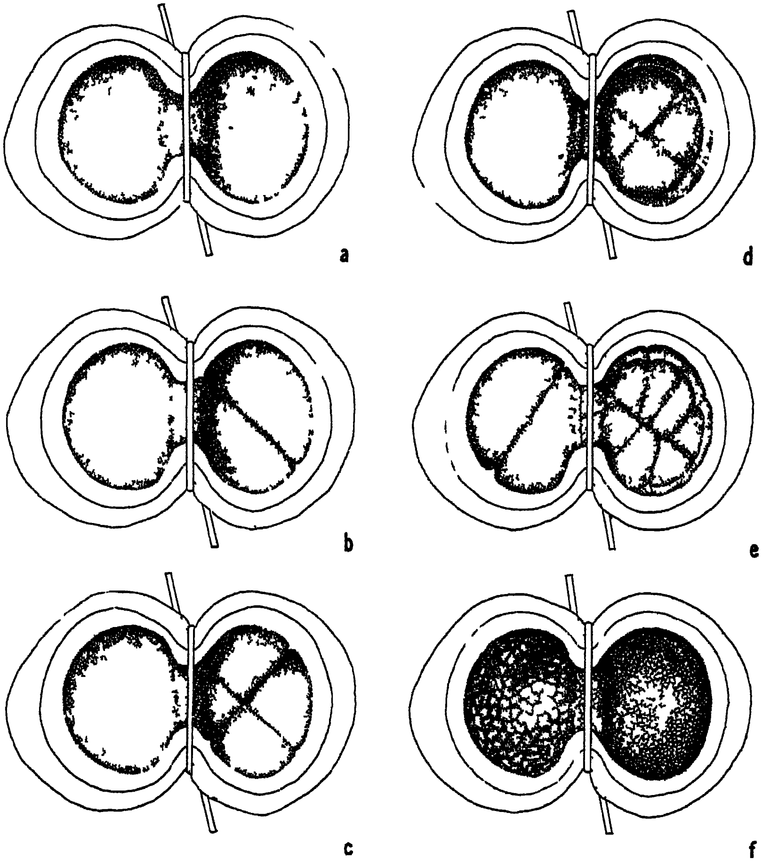


Figure 5 Drawings of cleavage of a partially constricted egg of *Triturus viridescens*, illustrating "delayed nucleation" of left-hand part Magnification about 14 times

a, immediately after constriction (about 20 minutes after insemination) Small black spot in center of light area, to right of hairloop, indicates position of second maturation spindle Larger dark spot above this area marks point of penetration of a spermatozoon

b, first cleavage of right half containing egg and sperm nuclei

c, second cleavage, one of the four cells connects with bridge, allowing its nucleus to enter

d, third cleavage, the nucleus in the bridge has divided and caused formation of a cleavage furrow to the left of center of bridge One of the daughter nuclei (*i e*, one of the eight nuclei of the eight-cell stage) has entered the left half

e, fourth cleavage, first division of left half

f, blastula stage, delay in beginning of cleavage of left half is clearly indicated by larger size of the cells

may develop because of the presence of supernumerary sperm nuclei. In rare cases, following constriction in the median plane, one may obtain twin larvae, one diploid, the other haploid. However, for our purposes, it would be advantageous to supply both halves with identical, diploid nuclei. Spemann (1914, 1928) showed that this may be done by partial constriction of the egg which leads to "delayed nucleation" of the originally non-nucleated half. The development of supernumerary sperm nuclei is usually inhibited in both halves of a partially constricted egg, as long as the connecting bridge remains wide enough (Fankhauser, 1925, 1930b).

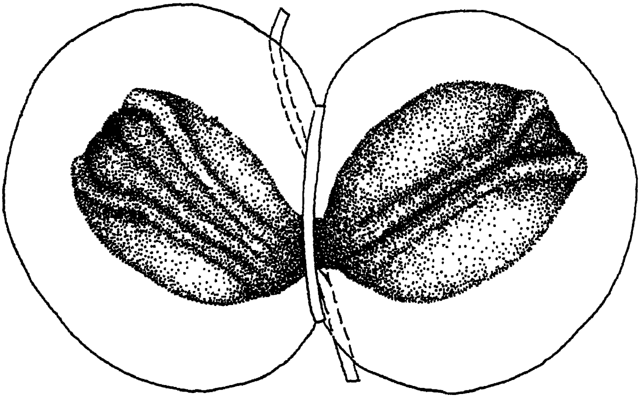


FIGURE 6. Neurula stage of egg that was partially constricted in future median plane, shortly after fertilization. Normal twin embryos. Difference in time of beginning of cleavage of the two halves is still clearly shown. Magnification about 25 times.

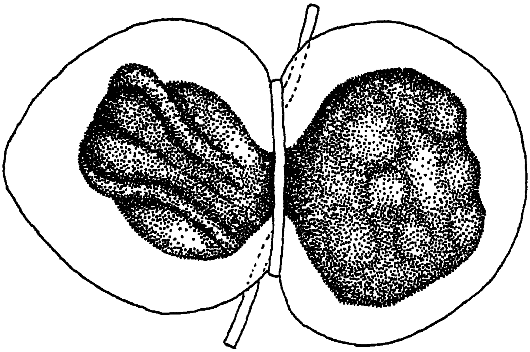


FIGURE 7. Neurula stage of another egg that was constricted in future frontal plane. Dorsal half develops normally, ventral half unable to form axial organs (right).

The experiment has been repeated many times with eggs of different species, *e.g.*, *Triturus viridescens* (FIGURE 5; *cf.* Fankhauser, 1932b). Cleavage begins in the half containing the fusion nucleus. Sooner or later, depending on the diameter of the connecting bridge, a descendant of the original nucleus will enter the bridge and divide there, as may be seen from the formation of a cleavage furrow. One of the daughter nuclei moves into the center of the uncleaved half and initiates a delayed cleavage which proceeds normally. As in experiments on the two-cell stage, one may obtain normal twins (FIGURE 6), or one normal and one ventral embryo (FIGURE 7). If the connection between the two parts of the egg is maintained during gastrulation, a double-headed monster may result

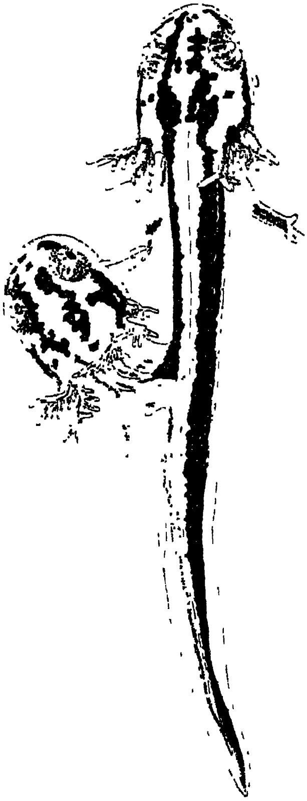


FIGURE 8. Two-headed larva developed from partially constricted egg, shown in neurula stage in FIGURE 6. 39 days old. Magnification about 14 times.

(FIGURE 8). If the bridge breaks, or is severed artificially by tightening of the hairloop, two normal, though slightly asymmetrical larvae may be produced (FIGURE 9).

Furthermore, following both complete and incomplete constriction of unsegmented eggs, a whole series of intergrades between ventral and normal embryos makes its appearance: embryos with weak axial organs, larvae with microcephaly of various degrees (FIGURES 10 and 11), larvae with strong unilateral defects limited to the head or trunk region (Fankhauser, 1930b, 1932b; Streett, 1940). This may be explained by the observation that the plane of constriction may form any angle with the invisible median plane of the egg so that the future center of organization is distributed between the two halves in various proportions. If eggs are constricted very lightly following fertilization, so that they are immobilized within the jelly capsule, the relation between the plane of constriction and the plane of symmetry may be determined at the time of appearance of the blastopore. Twenty-two eggs of *Triturus viridescens* tested in this way showed the following positions of the hairloop:

in median plane	3
small angle with median plane	4
oblique	8
small angle with frontal plane	3
in frontal plane	4

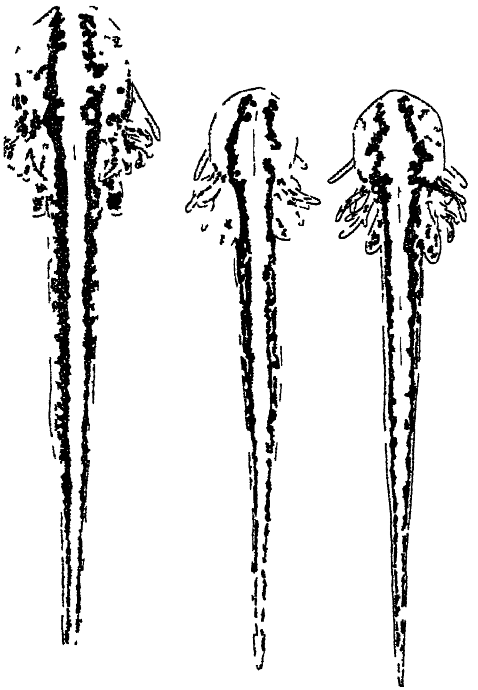


FIGURE 9 Control larva and twins produced by partial constriction of unsegmented egg in median plane The connection between the two halves of the egg broke after gastrulation Note smaller size of eye, balancer, and forelimb bud on right side of left-hand twin, smaller size of left forelimb bud of right-hand twin

On the basis of the results of experiments on the early gastrula stage, where the distribution of the dorsal lip area between the two halves may be observed directly, an egg fragment with a small lateral portion of the organization center would be expected to develop into a micro-axial or microcephalic embryo, while a fragment obtaining approximately one-half of the center would give rise to a "left" or a "right" larva with slight unilateral defects. The similarity of the types of abnormal embryos produced by halves of the unsegmented egg and of the gastrula raises the question to what extent regional differences in the future center of organization may be present at fertilization.

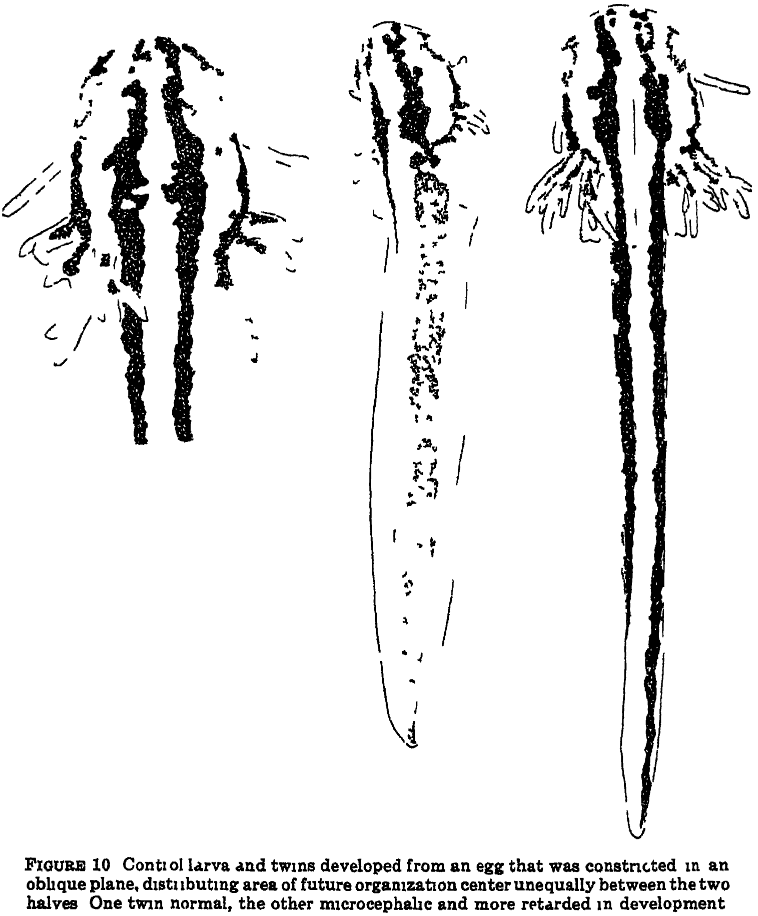


FIGURE 10 Control larva and twins developed from an egg that was constricted in an oblique plane, distributing area of future organization center unequally between the two halves One twin normal, the other microcephalic and more retarded in development

One more abnormality must be mentioned. Following median or nearly median constriction of the egg, the "left" twin, characterized by various degrees of underdevelopment of its right side, always has the heart and viscera in their normal position. Among the "right" twins, about 50 per cent show complete *situs inversus*. Reversal of asymmetry occurs whether the division of the egg takes place in cleavage or blastula stages (Spemann and Falkenberg, 1919), or shortly after fertilization (Fankhauser, 1930b). This indicates strongly that the typical asymmetry of the vertebrate body may be traced back to some slight difference in the invisible organization of the right and left halves of the egg at fertilization.

In the early days of experimental embryology, it was customary to make a sharp distinction between the eggs of Anurans and those of Urodeles. The former were supposed to show more "mosaic" organization, the latter to be examples of the "regulative" type. Although there undoubtedly exist differences in the degree of self-differentiation and of regulation between eggs of even closely related species of amphibians, there is no such clear-cut division between those of the two orders. Once the technical difficulties were overcome, Schmidt (1933), working in Spemann's laboratory, could show that constriction of eggs of *Bombinator* and of *Rana* produces essentially the same results as that of newts' eggs.

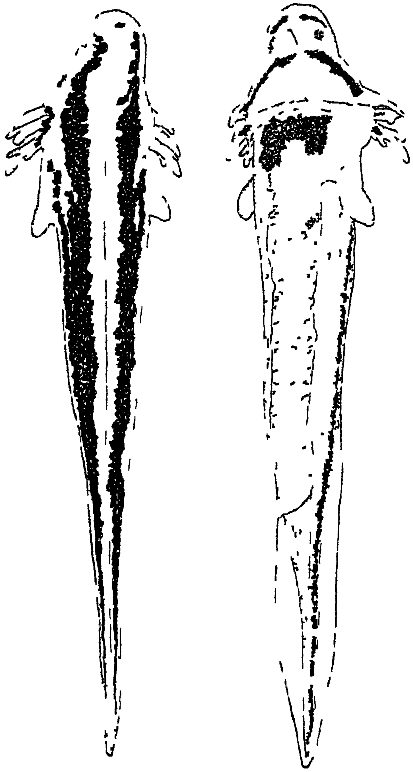


Figure 11. Dorsal and ventral views of a highly microcephalic larva developed from an egg fragment. The twin embryo was lost. Single balancer in ventral midline, between eyes.

While differences along the dorso-ventral axis are well established, those along the animal-vegetal axis have hardly been investigated. We need more information on the potencies of the isolated micro- and macromeres of the eight-cell stage, and of animal and vegetal fragments of unsegmented eggs and of cleavage stages, in extension of the work of Votquenne (1933), Vintemberger (1934, 1936), Streett (1940), and Stableford (1939). Such experiments are particularly important since tests of the blastula and gastrula stages have demonstrated a considerable capacity for self-differentiation on the part of the presumptive endoderm. Furthermore, there is the possibility of an early segregation of some material essential for the formation of the germ cells which, according to Bounoure (1934), takes place near the vegetal pole of the frog's egg before cleavage begins.

Analysis of the Nucleus

A discussion of the organization of the amphibian egg at the beginning of development would be very incomplete without a consideration of the nucleus and chromosomes. In recent years, geneticists have become increasingly interested in the cytoplasm of the egg, for several reasons. First, it is in the cytoplasm that most of the genes produce their effects. Second, presence of certain essential substances in the cytoplasmic substrate in limited amounts may explain some special phenomena, like competition between genes. Finally, the recent work of Sonneborn (1943) and Spiegelmann (1946) and others calls for the presence, in the cytoplasm, of self-reproducing units of nucleo-protein nature, probably derived from genes, which control the synthesis of proteins and enzymes. It is obvious that the embryologist, in turn, can contribute to the synthesis by paying more attention to the possibilities of experimenting with the chromosomes of the egg. I should like to mention briefly some lines of approach that may be followed:

a. Spemann's experiments on partial constriction demonstrated that a single nucleus of the eight- or sixteen-cell stage, which migrates into the non-nucleated half, is sufficient to initiate normal development of that half. Would a single nucleus of a more advanced stage, perhaps of an embryo, still be able to take the place of the normal fusion nucleus? Rostand (1943), repeating an experiment of Bogucki's, smeared unfertilized eggs of *Rana fusca* with embryo pulp obtained from a gastrula, then punctured them with a needle to induce parthenogenesis. Unsmeared control eggs gave the usual small number of dwarfed, unviable, haploid larvae. Smeared eggs produced similar tadpoles, but also a number of more vigorous larvae of normal size which survived for three to six weeks and were possibly diploid. Rostand suggests that, in these cases, a nucleus from one of the gastrula cells on the surface was dragged into the egg by the stylet and took over the functions of the cleavage nucleus, replacing the egg nucleus. It would be interesting to repeat the experiment on a larger scale and to work out the cytological details.

b. We know from observations on androgenetic salamander eggs and egg fragments that cell division may proceed in the complete absence of chromosomes. A blastula may be formed which includes large areas of non-nucleated cells containing actively dividing asters. Such eggs do not survive gastrulation (Fankhauser, 1934a; Fankhauser and Moore, 1941b). Thus, we know that the presence of chromosomes is necessary for normal development from this stage on. Other observations show that the requirements of the egg are more precise. From the time of gastrulation on, it needs at least one complete, balanced set of chromosomes. In the absence of the egg nucleus, sperm nuclei frequently form multipolar mitotic figures which result in abnormal distribution of chromosomes. With such varied, unbalanced sets of chromosomes present in different cells, development again comes to a standstill at the gastrula stage (Fankhauser, 1932b, 1934b; Kaylor, 1941).

On the other hand, complete sets of chromosomes may be subtracted or added with relatively slight effect on early development. Essentially normal development to early larval stages may take place in newts and other salamanders with anywhere from one to five sets of chromosomes, in spite of the fact that nuclear and cell size change in approximate proportion to the chromosome number (review in Fankhauser, 1945). Within a wide range of cell sizes, organogenesis is independent of the size of the individual building units.

c. The role of the chromosomes in development may also be studied by suitable crosses between various species of amphibians, some of which lead to an early arrest of development. As an example, we may mention the results of the genus cross between *Triton palmatus* and *Salamandra maculosa*, studied in Baltzer's laboratory by Schönmann (1938) and Lüthi (1938). Development invariably stops at the gastrula stage, when cells begin to die in various regions until the whole egg breaks up, although some cells still appear normal at this time. Signs of a lethal effect may be detected in the early blastula, when chromosomes frequently lag at anaphase and telophase. Obviously, there is some incompatibility between the *Salamandra* chromosomes and their *palmatus* environment which makes further development impossible. However, if a healthy piece of a hybrid gastrula is transplanted to a normal *Triton* gastrula, it may survive and form a harmonious part of various organs, such as the hindbrain and ear, with all the associated cartilages and muscles. Although other interpretations are possible, it seems likely that the cells of the transplant, although still normal at the time of the transfer to the normal host, were destined to die soon because of their hybrid constitution. Following transplantation, some factors residing in the healthy tissues of the host enabled the cells of the graft to overcome their inherent deficiency. Such stimulating effects of the adjacent normal host tissues have been demonstrated conclusively by Hadorn (1935, 1937) in chimaeras composed of parts of unviable merogonic hybrids and of normal embryos.

d. Time does not permit us to discuss in detail the experiments on hybrid androgenesis in which the maternal chromosomes are eliminated following cross-fertilization between two species. One should mention, at least, Hadorn's (1936) now classical experiment because it has a bearing on our interpretation of the cytoplasmic organization of the egg. In an embryo the cells of which contain cytoplasm of *Triton palmatus* and haploid nuclei of *Triton cristatus*, an incompatibility manifests itself soon after the closure of the neural folds, primarily in the head mesoderm. All embryos die at this stage (Baltzer, 1930). If a large piece of presumptive epidermis is taken from a still healthy embryo in the gastrula stage and transplanted to a normal gastrula of a third species, *Triton alpestris*, it will survive for months, until the host completes metamorphosis. At that time, specific characteristics appear in the epidermis. In *Triton cristatus*, the surface is smooth, while in *Triton palmatus* it shows toothlike projections formed by rows of flattened cells. The hybrid skin, with chromosomes of *Triton cristatus* and cytoplasm of *Triton palmatus*, shows the structure of the maternal species. This indicates that this characteristic was determined very early, while the egg was still in the ovary, under the influence of the maternal gene complex. The cytoplasm of the egg at fertilization would then contain some element which is responsible for the development of the characteristic skin pattern several months later. The recent work of Porter (1941) and of Moore (1946) on reciprocal hybridization and hybrid androgenesis between different local forms of *Rana pipiens* shows that other characteristics, appearing earlier in development, may also be influenced by the constitution of the cytoplasm of the egg.

Conclusions

The organization of the amphibian egg at fertilization, as it is known today, includes at least the following features:

(*1*) There must be a basis for the primary animal-vegetal polarity which may be connected, at least in part, with the distribution of yolk and associated substances rather than the properties of the egg surface alone, since the direction of animal-vegetal polarity may be completely inverted by centrifuging.

(*2*) There must also be a material basis for the bilateral symmetry of the embryo, *i.e.*, the dorso-ventral polarity, which is certainly present within a few minutes after fertilization, and may exist before. The direction of this axis may be shifted by secondary factors, such as a rotation imposed on the egg before the end of the first hour or, in some species, by the point at which the sperm enters the egg. There are indications that the dorso-ventral organization may be connected with the early localization of sulfhydryl and ribose-nucleic acid compounds derived from the germinal vesicle or from the cytoplasm surrounding it. These compounds, in turn, may be essential for the functioning of the future organization center, and for organogenesis in general.

(*3*) The typical asymmetry of the vertebrate body, which is expressed later in the position of the heart and viscera, also seems to be foreshadowed in the organization of the ooplasm. The nature of this factor is entirely unknown.

(*4*) The structure of the egg may be still more complex because of the presence of regional differences in the future organization center or in the presumptive endoderm.

(*5*) There are many indications of early interactions between cytoplasm and nucleus, for instance, the inhibition of the supernumerary sperm nuclei in the urodele egg, the normal coordination between chromosomal and centrosomal cycles of mitosis, and the peculiar reaction of the spermatozoa in immature eggs. All these adjustments are essential for the realization of normal cleavage.

(*6*) Slightly later, at the time of gastrulation, interactions between genes and cytoplasm assume importance, as is shown by the results of abnormal distribution of chromosomes through multipolar mitosis and by many hybridization experiments.

(*7*) Observations on reciprocal hybrids and on hybrid androgenesis show that various characteristics, appearing early or late in development, may be influenced by the constitution of the cytoplasm of the egg. The organization of the egg at fertilization may thus be complicated by the presence of components that are responsible for the appearance of such maternal characteristics.

(*8*) It is doubtful that we shall reach a real understanding of the already complex organization of the egg at fertilization unless we make a determined effort to trace its origins back into pre-fertilization stages, to the growing oöcyte in the ovary. During the weeks or months which the oöcyte passes in its follicle, the foundations of the future individual are laid down while all visible signs point to intense metabolic activity. Even the chromosomes take time out from their preoccupation with meiosis and undergo the profound and peculiar changes resulting in the still poorly understood "lampbrush" appearance, because they also are taking part in the synthesis of the elements of a new individual that will resemble the old. The young germ should be in an ideal condition to respond to some of the old questions of the embryologist. To be sure, it is less accessible to experimental procedures than it will be later on, but it may not prove to be as unapproachable as it is commonly supposed to be.

Bibliography

Ancel, P., & P. Vintemberger. 1933. Sur la soi-disant rotation de fécondation dans l'oeuf des amphibiens. C. R. Soc. Biol. 114: 1035–37.

1938a. Sur le déterminisme de la symétrie bilatérale dans l'oeuf de *Rana fusca*. Expérience permettant d'orienter le plan de symétrie bilatérale au gré de l'expérimentateur. C. R. Soc. Biol. 128: 95–97.

1938b. Sur le déterminisme de la symétrie bilatérale dans l'oeuf de *Rana fusca*. Rôle de la rotation d'orientation expérimentalement retardée dans les oeufs activés électriquement. C. R. Soc. Biol. 128: 412–414.

1938c. Sur le déterminisme de la symétrie bilatérale dans l'oeuf de *Rana fusca*. La localisation expérimentale du croissant gris est un phénomène reproductible au gré de l'expérimentateur pendant les deux premières phases de l'activation. C. R. Soc. Biol. 128: 414–416.

1938d. Sur le déterminisme de la symétrie bilatérale dans l'oeuf de *Rana fusca*. Orientation du plan de symétrie par le spermatozoide et création expérimentale d'un méridien de fécondation préférentiel. C. R. Soc. Biol. 128: 417–419.

1938e. Sur le déterminisme de la symétrie bilatérale dans l'oeuf de *Rana fusca*. Expériences dans lesquelles la rotation d'orientation et le spermatozoide agissent dans des sens différents. Conclusions générales. C. R. Soc. Biol. 128: 1212–14.

1938f. Sur le déterminisme de la symétrie bilatérale chez *Rana esculenta* et *Triton alpestris*. C. R. Soc. Biol. 129: 241–243.

Baltzer, F. 1930. Ueber die Entwicklung des Tritonmerogons *Triton taeniatus* × *cristatus*. Rev. Suisse Zool. 37: 325–332.

Banki, O. 1927. Die Lagebeziehung der Spermiumeintrittsstelle zur Medianebene und zur ersten Furche nach Versuchen mit örtlicher Vitalfärbung am Axolotlei. Verh. Anat. Ges. Kiel: 198–209.

1929. Die Entstehung der äusseren Zeichen der bilateralen Symmetrie am Axolotlei, nach Versuchen mit örtlicher Vitalfärbung. X^e Congrès Internat. de Zool. Budapest 1927, 1: 377–385.

Bataillon, E. 1929. Etudes cytologiques et expérimentales sur les oeufs immatures de batraciens. Arch. Entw.-Mech. 117: 146–178.

Bataillon, E., & Tchou Su. 1934. L'analyse expérimentale de la fécondation et sa définition par les processus cinétiques. Ann. Sci. Nat. (10me série) Zool. 17: 9–36.

Bounoure, L. 1924. Recherches sur la lignée germinale chez la grenouille rousse aux premiers stades de développement. Ann. Sci. Nat. Zool. 17: 67–248.

Brachet, A. 1910. La polyspermie expérimentale comme moyen d'analyse de la fécondation. Arch. Entw.-Mech. 30 (1): 261–303.

1912. La polyspermie expérimentale dans l'oeuf de *Rana fusca*. Arch. mikr. Anat. 79: 96–112.

1922. Recherches sur la fécondation prématurée de l'oeuf d'oursin (*Paracentrotus lividus*). Arch. Biol. 32: 205–248.

Brachet, Jean. 1944. Embryologie Chimique. Masson. Paris.

Dalcq, A. 1938. Form and causality in early development. Cambridge Univ. Press.

Dalcq, A., & J. Pasteels. 1937. Une conception nouvelle des bases physiologiques de la morphogénèse. Arch. Biol. 48: 669–710.

1938. Potential morphogénétique, régulation et "axial gradients" de Child. Mise au point des "bases physiologiques de la morphogénèse." Bull. Acad. Roy. Méd. Belg. (VI série) 3: 261–308.

Fankhauser, G. 1925. Analyse der physiologischen Polyspermie des *Triton*-Eies auf Grund von Schnürungsexperimenten. Arch. Entw.-Mech. 105: 501-580.

1930a. Zytologische Untersuchungen an geschnürten *Triton* Eiern. I. Die verzögerte Kernversorgung nach hantelförmiger Einschnürung des Eies. Arch. Entw.-Mech. 122: 117–139.

1932a. Cytological studies on egg fragments of the salamander *Triton*. II. The history of the supernumerary sperm nuclei in normal fertilization and cleavage of fragments containing the egg nucleus. J. Exp. Zool. 62: 185–235.

1932b. Cytoplasmic localization in the unsegmented egg of the newt, *Triturus viridescens*, as shown by the development of egg fragments. Anat. Rec. 54: 73.

1932c. The rôle of the chromosomes in the early development of merogonic embryos in *Triturus viridescens*. Anat. Rec. 54: 73.

1934a. Cytological studies on egg fragments of the salamander *Triton*. IV. The cleavage of egg fragments without the egg nucleus. J. Exp. Zool. 67: 349–393.

1934b. Cytological studies on egg fragments of the salamander *Triton*. V. Chromosome number and chromosome individuality in the cleavage mitoses of merogonic fragments. J. Exp. Zool. 68: 1–57.

1937. The development of fragments of the fertilized *Triton* egg with the egg nucleus alone ("gynomerogony"). J. Exp. Zool. 75: 413–469.

1945. The effects of changes in chromosome number on amphibian development. Quart. Rev. Biol. 20: 20–78.

Fankhauser, G., & C. Moore. 1941a. Cytological and experimental studies of polyspermy in the newt, *Triturus viridescens*. I. Normal fertilization. J. Morphol. 68: 347–385.

1941b. Cytological and experimental studies of polyspermy in the newt, *Triturus viridescens*. II. The behavior of the sperm nuclei in androgenetic eggs (in the absence of the egg nucleus). J. Morphol. 68: 387–423.

Hadorn, E. 1935. Chimärische Tritonlarven mit bastardmerogonischen und normalkernigen Teilstücken. Rev. Suisse Zool. 42: 417–426.

1936. Uebertragung von Artmerkmalen durch das entkernte Eiplasma beim merogonischen *Triton*-Bastard, *palmatus*-Plasma × *cristatus*-Kern. Verh. Dtsch. Zool. Ges. 97–104.

1937. Die entwicklungsphysiologische Auswirkung der disharmonischen Kern-Plasmakombination beim Bastardmerogon *Triton palmatus* (♀) × *Triton cristatus* ♂. Arch. Entw.-Mech. 136: 400–489.

Harrison, R. G. 1945. Relations of symmetry in the developing embryo. Trans. Conn. Acad. Arts & Sci. 36: 277–330.

Herlant, M. 1911. Recherches sur les oeufs di- et trispermiques de grenouille. Arch. Biol. 26: 103–835.

Kaylor, C. T. 1941. Studies on experimental haploidy in salamander larvae. II. Cytological studies on androgenetic eggs of *Triturus viridescens*. Biol. Bull. 81: 402–419.

Lehmann, F. E. 1942. Ueber die Struktur des Amphibieneies. Rev. Suisse Zool. 49: 223–228.

1945. Einführung in die physiologische Embryologie. Birkhäuser. Basel.

Lüthi, H. R. 1938. Die Differenzierungsleistungen von Transplantaten der letalen Bastardkombination *Triton* ♀ × *Salamandra* ♂. Arch. Entw.-Mech. 138: 423–450.

Mayer, B. 1939. Versuche zum Nachweis der Induktionsfähigkeit jüngster Entwicklungsstadien von *Triton*. Naturwiss. 27: 277.

Moore, J. A. 1946. Hybridization between *Rana palustris* and different geographical forms of *Rana pipiens*. Proc. Nat. Acad. Sci. 32: 209–212.

Needham, J. 1942. *Biochemistry and Morphogenesis*. Cambridge Univ. Press.

Pasteels, J. 1937. Sur l'origine de la symétrie bilatérale des amphibiens anoures. Arch. Anat. Micr. 33: 279–300.

1938. A propos du déterminisme de la symétrie bilatérale chez les amphibiens anoures. Conditions qui provoquent l'apparition du croissant gris. C. R. Soc. Biol. 129: 59–61.

1940a. Recherches sur les facteurs initiaux de la morphogénèse chez les amphibiens anoures. III. Effets de la rotation de 135° sur l'oeuf insegmenté, muni de son croissant gris. Arch. Biol. 51: 103–149.

1940b. Recherches sur les facteurs initiaux de la morphogénèse chez les amphibiens anoures. IV. Centrifugation axiale de l'oeuf fécondé et insegmenté. Arch. Biol. 51: 335–386.

1941. Recherches sur les facteurs initiaux de la morphogénèse chez les amphibiens anoures. V. Les effets de la pesanteur sur l'oeuf de *Rana fusca* maintenu en position anormale avant la formation du croissant gris. Arch. Biol. 52: 321–339.

Porter, K. R. 1941. Diploid and androgenetic haploid hybridization between two forms of *Rana pipiens*, Schreber. Biol. Bull. 80: 238–264.

Rostand, J. 1943. Essai d'inoculation de noyaux embryonnaires dans l'oeuf vierge de grenouille. La Revue Scientifique 81: 454–456.

Rotmann, E. 1943. Entwicklungsphysiologie. Fortschr. Zool. 7.

Ruud, G. 1925. Die Entwicklung isolierter Keimfragmente frühester Stadien von *Triton taeniatus*. Arch. Entw.-Mech. 105: 209–293.

Ruud, G., & H. Spemann. 1922. Die Entwicklung isolierter dorsaler und lateraler Gastrulahälften von *Triton taeniatus* und *alpestris*, ihre Regulation und Postgeneration. Arch. Entw.-Mech. 52: 95–166.

Schmidt, G. A. 1933. Schnürungs- und Durchschneidungsversuche am Anurenkeim. Arch. Entw.-Mech. 129: 1–44.

Schönmann, W. 1938. Der diploide Bastard *Triton palmatus* ♀ × *Salamandra* ♂. Arch. Entw.-Mech. 138: 345–375.

Smith, B. G. 1922. The origin of bilateral symmetry in the embryo of *Cryptobranchus alleghenienisis*. J. Morphol. 36: 357–399.

Sonneborn, T. M. 1943. Gene and cytoplasm. Proc. Nat. Acad. Sci. 29: 329–343.

Spemann, H. 1901. Entwicklungsphysiologische Studien am *Triton*-Ei. Arch. Entw.-Mech. 12: 224–264.

1902. Entwicklungsphysiologische Studien am *Triton*-Ei. II. Arch. Entw.-Mech. 15: 447–534.

1914. Ueber verzögerte Kernversorgung von Keimteilen. Verh. Deutsche Zool. Ges., 24. Jahresvers. Freiburg i. Br. 216–221.

1928. Die Entwicklung seitlicher und dorsoventraler Keimhälften bei verzögerter Kernversorgung. Z. wiss. Zool. 132: 105–134.

Spemann, H., & H. Falkenberg. 1919. Ueber asymmetrische Entwicklung und *situs inversus viscerum* bei Zwillingen und Doppelbildungen. Arch. Entw.-Mech. 45: 371–422.

Spiegelman, S., & M. D. Kamen. 1946. Genes and nucleoproteins in the synthesis of enzymes. Science 104: 581–584.

Stableford, T. 1939. The potency of the isolated vegetal hemisphere (presumptive endoderm) of the blastula of *Amblystoma punctatum*. Anat. Rec. 75: 35–36.

Streett, J. C. 1940. Experiments on the organization of the unsegmented egg of *Triturus pyrrhogaster*. J. Exp. Zool. 85: 383–408.

Vintemberger, P. 1934. Résultats de l'autodifférentiation des quatre macromères isolés au stade de huit blastomères, dans l'oeuf d'un amphibien anoure. C. R. Soc. Biol. 117: 693–696.

1936. Sur le développement comparé des micromères de l'oeuf de *Rana fusca* divisé en huit: *a*) après isolement, *b*) après transplantation sur un socle de cellules vitellines. C. R. Soc. Biol. 122: 927–930.

Vogt, W. 1929. Die Beziehungen zwischen Furchung, Hauptachsen des Embryo und Ausgangsstruktur im Amphibien-embryo, nach Versuchen mit örtlicher Vitalfärbung. Sitz.-Ber. Ges. Morph. & Phys. München 37: 60–70.

Votquenne, M. 1933. Expériences de destruction des micromères dorsaux de l'oeuf de *Rana fusca* au stade VIII, et interprétation des résultats par la méthode des colorisations vitales localisées. Arch. Biol. 45: 79–154.

Discussion of the Paper

Dr. Robert Chambers (*New York University, New York, N. Y.*):

I wish to continue the point I raised in the discussion of Dr. Costello's paper in which I suggested that the sperm or monaster in many marine eggs may be concerned with the reorientation of formative stuffs. The monaster in the larger and more highly lecithal amphibian egg does not develop to the extent that it does in the echinoderm egg in which the entire egg substance is involved. Is it possible that the lack of a sufficiently extensive aster permits dislocation in these large eggs by fragmentation, or by abnormal sedimentation, or centrifugation to upset the organization pattern?

Dr. L. T. Stableford (*Department of Biology, Lafayette College, Easton, Pennsylvania*):

There is evidence from the cytological observations of Schultze, Hibbard, and others that the polarity of the amphibian egg is determined at an early stage of the oöcyte since vitellogenesis is initiated excentrically in that portion of the oöcyte which later becomes the vegetal hemisphere of the egg. With this in mind, I would suggest that the relation between yolk and cytoplasm in the egg is more than "gravitational," that it is, rather, intimately structural.

This concept of a yolk-cytoplasm complex in the vegetal hemisphere runs contrary to the conclusions commonly drawn from inversion and centrifugation studies. It is usually assumed that, when the egg is inverted, the yolk spherules change position under the influence of gravity and mechanically force a reorganization of the egg. It should be pointed out, however, that this experiment involves, in addition to gravity, the effect of mechanical pressure since the egg must be compressed between glass plates or in a tube to invert it. Granting that Pflüger showed pressure itself to be without effect on the development of the egg, it seems reasonable to suggest that the combination of pressure and gravity disturbs the structure of the egg so radically that the shifted material is actually a yolk-cytoplasm complex.

Further evidence for a yolk-cytoplasm complex as a structural component of the egg comes from the work of Todd in 1940. He indicated that, when the amphibian egg is centrifuged, there is first a movement of liquid material to the animal pole of the egg and that this material flows between the large yolk spherules; it is apparently only when the centrifugal force is strong that the characteristic packing of yolk at the vegetal pole occurs.

I bring up this idea of a possible yolk-cytoplasm complex mainly to call attention to the internal portions of the egg since interest seems to be centering almost exclusively on the cortex or surface coat at present.

DR. G. FANKHAUSER:

In reply to Dr. Chambers, let me say that observations on eggs of several species of amphibians have shown that the fully developed sperm aster is very large and may extend through the entire animal hemisphere. However, the maximum size is not reached until two and a half hours after fertilization (in eggs of newts). While the growth of the sperm aster thus may be an important factor in the rearrangement of certain egg materials, there is no direct evidence that it is instrumental in the segregation of developmental potencies in different parts of the egg.

The views expressed by Dr. Stableford are in general agreement with those developed by Dalcq and Pasteels, who emphasize the active role of the "yolk" in early development. Future investigations of this problem should include a more exact determination of the roles played by the material of the yolk platelets themselves and by other substances that might be closely associated with these bodies.

SIGNIFICANCE OF THE CELL MEMBRANE IN EMBRYONIC PROCESSES

BY JOHANNES HOLTFRETER*

Biology Department, University of Rochester, Rochester, N. Y.

EVIDENCE of the embryological significance of the cell membrane may be derived from a consideration of this structure in connection with the following phenomena: cellular permeability, amoeboid motility, morphogenetic movements, cytoplasmic division, selective adhesiveness, cell polarity, and cellular differentiation. It is clear that such a wide range of phenomena cannot be discussed exhaustively within the frame of the present review. The aim of this paper is, therefore, to give not more than a tentative synopsis of the embryological functions of the cell membrane and to discuss the physico-chemical and physiological properties of this structure only as they may serve this purpose. This paper will mainly be concerned with observations on amphibian development, although it is realized that some of the phenomena to be touched upon have been more thoroughly investigated in other forms, such as the echinoderms. The discussion will include numerous unpublished observations.

Functions of the Coat in the Amphibian Egg

Physico-Chemical Properties of the Coat. The array of protective envelopes investing a fertilized amphibian egg resembles that described for the sea urchin egg by Chambers (1938, 1940). Apart from the external gelatinous layers and the vitelline membrane, the amphibian egg and its subsequent stages are covered by a film, or coat, which is firmly attached to the underlying cell membrane and which seems to be comparable to the hyaline layer in echinoderm eggs. The integrity of the coat requires the presence of small amounts of calcium in the immersion fluid (Holtfreter, 1943a). Amphibian embryos which are placed in an isotonic solution lacking calcium ions fall apart into single cells, and a mucilaginous substance passes into solution. Similar mucus formation and disaggregation occur when the pH of the balanced salt solution is raised above 9.6 or lowered below 4.2, or when solutions of sodium citrate or oxalate are applied. As in the case of disintegrated echinoderm blastomeres, the amphibian cells reaggregate and survive the treatment if, within a restricted period, calcium is restored to the external medium. A new coating substance may be secreted by the outermost surface of the ectodermal and endodermal cells.

---

* In dedicating this paper to Professor K. von Frisch, I am expressing my great admiration for his scientific work, and my indebtedness to him for the inspiration and kind assistance I received while I was a member of his former Zoological Institute at Munich.

In sea urchin eggs, the hyaline coat evidently consists of proteins. Excess of calcium in the sea water renders this layer rigid and insoluble, while acidulated sea water (pH 3.5) makes it disappear (Moore, 1928). The hyaline layer is furthermore attacked by trypsin, and it is negatively birefringent in the radial direction, indicating the presence of tangentially oriented protein micells (Runnström, Monné, and Broman, 1943). Based upon the data on the susceptibility of the coat in amphibian embryos to acids, bases, and calcium-free solutions, it may be assumed that here, too, proteins constitute the essential components of this structure. Alcohol and other protein coagulants render the coat brittle.

Dissection experiments on living amphibian embryos show that the coat is apparently not a living and indispensable part of the egg and of the epithelia deriving from the egg surface. While covering and interconnecting the peripheral cells as a syncytial layer, the coat, where it forms intercellular bridges, can be drawn out between glass needles into long contractile threads which may be cut off without affecting in any way the viability of the attached cell. The surface of the coat is normally semi-solid and non-adhesive, capable of forming wrinkles when stretched, and being in the living egg in a state of elastic tension. Because of these properties, the coat represents the essential structure which unites the blastomeres into a closely packed body.

As compared with the proximal, uncoated side of the surface epithelia, the coated cell surface is considerably less permeable to water, electrolytes, and various toxic substances. Whereas whole embryos, or embryonic fragments, which are entirely covered by a layer of coated cells can be cultured for long periods in strongly hypotonic media containing traces of calcium, uncoated fragments cytolyze in it within a short time. Uncoated cells also show an enhanced susceptibility to hypertonic solutions and vital dyes. It is this reduced permeability of the surface layer which enables the amphibian egg to develop normally in tap water.

A protective coat does not seem to exist in the ovarian eggs of the frog, since isolated eggs of this stage die rapidly in various solutions which do not impair the viability of mature eggs. On the other hand, overripeness is associated with a softening of the egg surface, which is reflected in its reduced resistance to centrifugation and its reduced capacity of closing wounds. It is not unlikely that the abnormal cleavage and gastrulation patterns observed in fertilized overripe frog's eggs (Witschi, 1930; Briggs, 1941) are predominantly due to a partial disintegration of the egg surface, involving probably both the coat and the underlying plasma membrane.

Susceptibility to Mechanical Agitation. This double layer disintegrates readily when disturbed mechanically. The susceptibility can be demonstrated by placing a number of fertilized frog's eggs, with all their coverings intact, in a bag of cheesecloth and letting a quick succession of water drops hit the external jelly of the eggs. If the drops fall from a height of not more than about 10 cm., the eggs suddenly swell strongly and disintegrate within an hour. It is surprising that the eggs succumb to the rhythmic percussions while they are still enclosed in their protective envelopes and without being visibly deformed by the impact of the falling drops. The observation recalls the early experiments of Dareste (1891) which have recently been taken up by Landauer and Baumann (1943), and which showed that shaking of unincubated chicken eggs tends to produce drastic developmental abnormalities. Fauré-Fremiet (1932) and Battle (1948) found that the eggs of various fishes become rapidly liquefied or develop abnormally when mechanically agitated. Similar liquefying effects of shaking or other mechanical irritations have been observed in Protozoa, bacteria, erythrocytes and other cells (Koelsch, 1902; Angerer, 1936; Chambers, 1924). Frequently, death of the cells is preceded by the formation of hyaline blisters and by considerable swelling, which suggests that disintegration is initiated by a molecular disarrangement of the surface membrane, producing at first increased permeability, and subsequently breakdown (see p. 731).

Significance of the Coat in the Process of Gastrulation. When the cells of an amphibian gastrula enter into the phase of morphogenetic movements, it is in the first place the investing coat which integrates the amoeboid activity of the individual cells into the coordinated and synchronized movements of entire germ layers (Holtfreter, 1943b, 1944). By virtue of this superficial elastic sheet which cements the peripheral cells together into a continuous layer, gliding movements of the cells in any region of the gastrula are transmitted to adjacent regions. Similar effects of a tangential pull can be observed in a layer of ectoderm which is proceeding to close in on a wound inflicted to this layer. It should be emphasized, however, that the motive forces for the gastrulation movements of spreading and invagination cannot be ascribed to the properties of the coat, but originate within the living cells proper. Embryonic cells which are not held together by a syncytial coat are perfectly capable of spreading over an organic or inorganic substratum, and they will slip, singly or in groups, into the depth of a layer of endoderm with which they have been brought into contact. The coat merely regiments these amoeboid movements so that they become collective events. Actually, the tensile strength and contractility of this structure counteract the movements of invagination. Cells which migrate singly into the interior of a gastrula can do so only by detaching themselves from the coat. Where such a detachment does not occur, the invaginating cells become stretched into cylindrical or even filiform bodies, which, through their combined efforts, draw the coated surface inward in the form of an archenteron or other cavities.

It follows from the above considerations that agents which reduce the tensile strength of the coat, or remove the layer altogether, will have drastic effects on the morphogenetic movements, although invagination will not necessarily be suppressed. Agents having such a dispersive or liquefying effect on the coat have already been mentioned; they are salt solutions which are hypertonic, or free of calcium, or have a pH above 9.6 or below 4.4. The suppressive effect of lithium on the morphogenetic movements in amphibian embryos does not seem to differ markedly from that of other monovalent cations.

A weakening of the coat results in exogastrulation. The movements most inhibited are the epiboly of the ectoderm and the involution of the endoderm, whereas the mesodermal tissues freely invaginate into the interior of the endoderm (Holtfreter, 1933). Ectoderm, when its coat has been removed or has lost its original non-adhesiveness, may even become embedded within the endoderm. Morgan (1903), by applying hypertonic solutions of lithium chloride, obtained such a partial inversion of the germ layers in frog embryos. In a similar way, by making use of the coat-dissolving action of sodium citrate (hypo- to hypertonic solutions), the writer was able to produce frog gastrulae in which the entire "decoated" ectoderm, together with the mesoderm, sank into the interior, while the endoderm became the external layer of the whole embryo.

Exogastrulation can be brought about either by culturing pre-gastrula stages permanently in solutions having only a softening effect upon the coat, or by subjecting the early gastrula for a brief period to media, such as alkali or sodium citrate, which produce an immediate and total breakdown of the coat, then returning them to physiologically normal conditions. The idea of Jenkinson (1906) that hypertonicity of the external medium is not the sole agent which causes exogastrulation, is strongly supported by these experiments where the agents were applied in hypotonic solutions, and where the treatment lasted not longer than 15 to 20 minutes. This gave the disaggregated cells full opportunity to reestablish the continuity of the embryo and to continue differentiating.

It would be unwise to attempt an explanation of the whole exogastrulation syndrome in terms of a weakening or dispersal of the coating substance. However, it is reasonable to assume that the integrity of the coat is at least one of the most important factors safeguarding the normal display of the movements of gastrulation and neurulation.

Fate of the Coat in Later Stages of Development. In consequence of the infolding of all except the prospective epidermal areas of the egg surface, the coat is partly carried into the interior of the embryo and there transformed into the inner linings of the intestinal epithelium, of the neural tube, and possibly of the kidney tubules. In this new position, the surface layer appears to retain its original non-adhesiveness, thereby preventing a fusion of the tubular walls at places where they touch each other. On the other hand, the somite-notochord material which originally forms part of the egg surface seems to lose the coating substance, since after having invaginated it breaks up into tissues which are entirely adhesive. Pieces of gastrula ectoderm, when grafted into the mesenchyme of older hosts, may either retain their epithelial continuity forming cysts, where the coated surface is turned inside, or they, too, break up into cell strands which have no apparent proximo-distal polarity, are adhesive, and disperse within the host tissues.

During the course of development, the coated surface of the epidermis undergoes a progressive hardening, a process which can be demonstrated by subjecting tissues of different developmental stages to the tension at the interface between air and aqueous media having different surface tensions. Such experiments show that the resistance of the coated ectoderm to being torn apart by the interfacial forces increases considerably after the end of neurulation, and that already in younger stages the cohesion of the coated surface is markedly stronger than that of the inner, uncoated surface of the epithelium (Holtfreter, 1943b). Ectoderm cells which are isolated singly in physiological salt solution acquire, within a few days, an entirely rigid and non-adhesive surface. At the same time, their amoeboid motility disappears, while their ciliary movements continue. This solidifying effect of the external medium is further exemplified in those ectoderm cells which, in a normal embryo, move from deeper layers into the outer surface epithelium, where they acquire the characteristics of coated cells.

It seems reasonable to suppose that the progressive hardening of the epidermal surface is largely a result of the progressive formation of insoluble calcium proteinates, since the length of time required to disintegrate the surface coat in a given concentration of potassium oxalate, of sodium citrate, or of sodium chloride, increases with the developmental stage of the embryo. Swimming larvae can no longer be disintegrated by these methods, indicating an insoluble and possibly denatured state of the surface proteins.

Observations on the Properties and Functions of the Cell Membrane

The eggs of various animals, for instance those of some Nematodes (Spek, 1918) and of the trout (Yamamoto, 1940), before or after the onset of the cleavage process, exhibit undulating or rotating movements of their surface. In the oocyte stage, some eggs, such as those of echinoderms, may form pseudopodia of various shapes, particularly under the influence of penetrating spermatozoa, or of chemicals with parthenogenetic properties (Seifriz, 1927; Runnström, 1928; Runnström and Monné, 1945; Harvey, 1938). Intact amphibian eggs do not seem to perform surface movements although, under abnormal conditions, an unfertilized frog's egg may form local bulges and papillae which tend to become pinched off by constriction (Holtfreter, 1946a). This inertness of the amphibian egg may be the result of mechanical restraint exerted by the relatively tough coat. On the other hand, the uncoated side of the blastomeres, facing the interior of a morula or blastula, is normally endowed with filiform or knob-shaped pseudopodia (Holtfreter, 1943b). With progressive differentiation, cell motility increases, as may be seen by the morphogenetic movements and by the subsequent migrations and amoeboid changes of the cells. The following data seem to indicate that it is the external cell membrane rather than intracellular structures which are responsible for these amoeboid movements (Holtfreter, 1946–1947).

Structural Organization of Embryonic Amphibian Cells. The architecture of any cell from early amphibian embryos closely resembles the general organization found in *Amoeba proteus* (Chambers, 1924; Mast, 1926) and in *Pelomyxa* (Wilber, 1946). There are four concentrically arranged layers: an inner core of semiliquid "plasmasol," containing the nucleus, the structural cytoplasm, and all of the granular inclusions. Even in a resting cell, Brownian movement causes damped translocation of the granules. This material is enclosed in a wall of "plasmagel," the viscous consistency of which prevents dislocations of particles embedded in it. The third layer consists of the clear "ectoplasmic fluid" which occasionally contains a few granules derived from the granulated endoplasm. The unrestricted translocations which the granules exhibit when entering this layer indicate clearly its highly fluid consistency. The hyaline fluid is surrounded by a well delineated plasmalemma, or cell membrane, which may fold into wrinkles, suggesting a semi-solid consistency.

The thickness of the hyaline layer varies with the composition of the external medium. Under approximately normal physiological conditions, the fluid accumulates only in regions where the cell surface bulges out into a pseudopodium, and it fades out from view where the cell membrane becomes temporarily apposed to the plasmagel (FIGURE 1). However, although not visible everywhere, this layer seems to form a continuous shell around the endoplasmic granuloplasm. The continuity can be readily demonstrated when the amount of ectoplasmic fluid is increased following immersion of the cell in isotonic saline solutions lacking calcium, or having a pH between 9 and 10. Solutions more acid, and rich

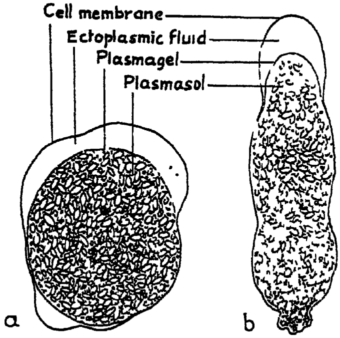


FIGURE 1. Protoplasmic structure of embryonic amphibian cells.

in calcium, have the opposite effect of reducing the hyaline layer. This agrees with observations on amoebae (Pantin, 1923, 1926) and on amoebocytes of *Limulus* (Loeb, 1928).

The cells constituting the surface of the amphibian gastrula possess a proximo-distal polarity which seems an expression of an inside-outside gradient of the egg. When cells from the periphery of the embryo are isolated mechanically, or by means of coat-dissolving chemicals, the former coated side becomes the posterior cell pole which is characterized by a comparatively stronger contractile power of its surface membrane, by the accumulation of pigment, and by the fact that this side has a reduced tendency to form hyaline protrusions. The uncoated proximal side of the cell becomes the region of pronounced amoeboid activity. It represents, therefore, the advancing anterior pole when the cell is resorting to locomotion. Isolated embryonic cells of any germ layer, but especially those of the neural plate, tend to stretch themselves along their proximo-distal axis into cylindrical bodies having an anterior cap filled with ectoplasmic fluid, while the posterior pole is usually marked as a blackish knob showing surface wrinkles. Following exposure of the cylindrical cell to liquefying agents, such as alkali or sodium citrate, hyaline protrusions also appear along the side walls suggesting that there, too, the cell membrane has remained separated from the plasmagel by a thin layer of fluid (FIGURE 2). With further uptake of water, the ectoplasmic fluid forms a spacious shell around the entire endoplasm, followed by a rounding-up of the whole cell into a more or less spherical body.

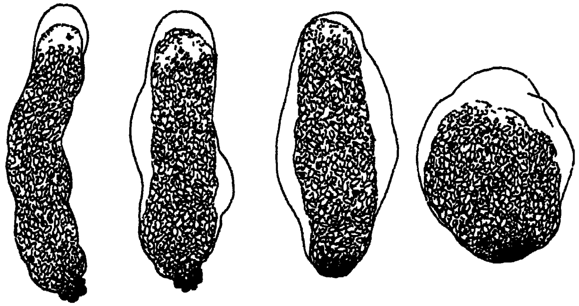


FIGURE 2. Exposure to alkali of a cylindrical cell from the medullary plate causes a lifting-off of the cell membrane from the plasmagel tube and a rounding up of the cell.

Amoeboid Movements. According to the widely accepted theory of Mast (1926, 1941), form changes and locomotion in *Amoeba* result from contractions of the posterior portion of the plasmagel tube, which push the enclosed plasmasol forward against the plasmagel wall bordering at the hyaline cap of a pseudopod. Subsequently, the frontal gel barrier becomes solated, breaks down, and the endoplasm surges into the fluid of the cap. There, the granuloplasm is presently regelated at its surface, thus forming an anterior extension of the plasmagel tube. The plasmalemma of the hyaline cap is assumed to be pushed out passively by the hydrostatic pressure of the forward-streaming axial protoplasm. Repetitions of this cycle would result in a continuous transport of the endoplasm from the posterior into the anterior cell portion, *i.e.*, a shifting forward of the whole cell body. Thus, the essential mechanism of locomotion is supposed to consist of localized and reversible sol-gel formations of the granulated endoplasm, with the cell membrane passively yielding to the local variations of the internal turgor.

In elaborating upon this theory, Lewis (1942) assumed that a conversion of the plasmasol into a more viscous state would automatically result in a contraction of this material. Contractions occurred not only in the posterior region of the cell, but Lewis (1933, 1942) observed in lymphocytes and other vertebrate cells that constriction rings passed in regular intervals over the entire cell surface, traveling in an anteroposterior direction. This peristaltic activity was assumed to reside in the plasmagel tube. Lewis concluded that the translocations of the inner protoplasm resulting from these constrictions, together with localized cyclical sol-gel formations, are the essential mechanism of cellular locomotion.

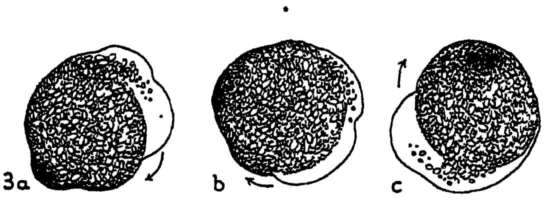
The processes described by Mast and Lewis may occur, likewise, in isolated cells from different stages and germ layers of amphibian embryos. However, the forward-streaming and cyclical sol-gel conversions of the endoplasm observed here, appear to be the consequence rather than the cause of amoeboid movements. Form changes and locomotion of embryonic cells are not necessarily associated with sol-gel transformations of the protoplasm, but seem to be primarily due to alternate expansions and contractions of localized regions of the cell membrane. This conclusion is based upon the following data.

Cellular Motility Associated with Endoplasmic Sol-Gel Formations. Experimental variations of the composition of the culture medium show that both the state of viscosity of the endoplasm and the motility of the cell membrane change with the external conditions, but these two phenomena are not strictly correlated with each other. In a balanced isotonic salt solution ("standard solution") with pH between about 7 and 8.5, isolated cells from the different germ layers of a gastrula or neurula may exhibit the following kinetic phenomena.

Rotating Lobopodia in Spherical Cells. At one or several points of the spherical cell the outer membrane is lifted off the plasmagel and forms hyaline bulges which tend to move around the circumference of the cell, but usually avoid the posterior cell pole. After a bulge has attained maximal size, the underlying portion of the plasmagel may liquefy, and the endoplasmic particles surge into the ectoplasmic fluid where they are freely dispersed and thrown about in rapid motion. Subsequently, the liberated particles undergo a closer packing, become almost immobile, and are refurnished with a capsular plasmagel wall lying closely beneath the cell membrane. Meanwhile, an adjacent region of the cell surface has been growing into a hyaline lobopodium, and the process of solation and subsequent regelation of the inundated portion of the endoplasmic capsule is repeated. Rotating lobopods of a similar kind have been observed in amoebae (Rhumbler, 1898; Jennings, 1904; Pantin, 1923) and in the amoebocytes of invertebrates (Loeb, 1928).

If the embryonic cells are suspended in isotonic saline lacking calcium, or having a pH above 9, they swell considerably through increase of ectoplasmic fluid. A large hyaline cap is formed which rotates rapidly around the surface, while the endoplasmic capsule which is temporarily overflowed by this wave, reacts after a latent period with rapid and merely superficial solations and regelations (FIGURE 3). The belatedness of the endoplasmic eruptions makes it very improbable that this process is the cause of the outbulging and of the rotation of the hyaline lobopod. On the contrary, it seems that the local solations of the plasmagel are of a secondary nature, caused by a liquefying action of the supernatant fluid, the volume of which remains constant throughout the successive cycles of movement. That the cell surface moves independently of the endoplasmic eruptions is borne out by the fact that the larger lobopods consist of several bulges which are separated by notches of contraction, traveling wave-like around the circumference of the cell.

Peristaltic Movements and Locomotion in Cylindrical Cells. The main difference between the pattern of movement in a spherical and a cylindrical cell is that in the latter the hyaline cap and the sol-gel process are confined to the very end of the anterior region. Instead of showing rotating notches of constriction, the elongated cell possesses constricting rings which travel in fairly regular intervals from the anterior to the posterior pole (FIGURE 4). In addition, the elongated cell may perform bending and twisting motions along its main axis and may undergo longitudinal elongations and contractions of the anterior body portion, while the posterior portion is periodically contracted. This pattern of vermiform movements can be exhibited by isolated cells of various germ



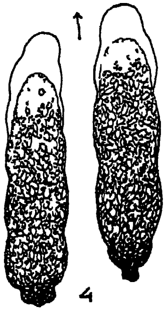


FIGURE 3. Rotating movement and endoplasmic eruptions in a cell isolated from gastrula ectoderm.

FIGURE 4. Peristaltic constrictions and locomotion in a cell from the medullary plate.

layers, although it is most pronounced in the strongly elongated cells from the neural plate.

Since the waves of constriction first appear within the purely hyaline cap, it is clear that they cannot result from a kinetic activity of the plasmagel tube but must be due to an autonomous contractility of the cell membrane. The endoplasmic tube which is, in general, closely associated with the cell membrane seems to be passively molded by the constrictions passing over the cell surface. When a constriction ring approaches the posterior cell pole, the strength of contraction increases, leading to the temporary formation of a small tail knob having a wrinkled surface. The anterior cap never becomes wrinkled but is periodically pushed out in the form of conical bulges. It is these alternating longitudinal elongations and contractions of the anterior region, combined with the periodical shortenings of the tail end which, in the presence of a supporting surface of friction, will produce a shifting forward of the whole cell body. The mechanism of locomotion is thus comparable with that of a creeping earthworm. It is obvious that the peristaltic constrictions, which both in the cell and in the worm occur together with the longitudinal extensions and contractions, cannot be considered as effective movements of locomotion.

It may be mentioned already at this point that the cycle of surface movements may go on indefinitely in the total absence of endoplasmic streamings. Sometimes, however, when a sufficiently large amount of fluid has accumulated in the hyaline cap, the frontal wall of the plasmagel liquefies and the endoplasm flows into the cap where it is presently regelated at its surface. Vital staining of the posterior cell portion shows that, in contrast to the conditions in *Amoeba*, there is no real axial streaming of the plasmasol, but that it is more or less the same circumscribed portion of the anterior endoplasm which is periodically released into the ectoplasmic fluid. This restriction of the gel-sol process to the very tip of the pseudopod is still more conspicuous in the extended processes of more highly differentiated cells.

Cellular Motility in the Absence of a Sol-Gel Cycle. The gelated state of the endoplasm, and hence the frequency of breaks in the plasmagel wall, depend upon the concentration of calcium ions in the culture fluid. Cells which are kept in isotonic solutions of the chlorides of Na, K, or Li, develop a very broad hyaline space, into which the endoplasmic core is almost uninterruptedly releasing its granuloplasm. Under these conditions, regelation of the liberated endoplasm is considerably delayed. It is completely inhibited if sodium citrate or oxalate is applied, either in addition to the chlorides or in pure solutions. These calcium-antagonizing substances inhibit gelation both in hypertonic and hypotonic concentrations. In the former case (0.5 to 2 per cent solutions) the cells shrink and the hyaline space is reduced to a small blister. The endoplasm underlying the blister lacks the smooth surface of a plasmagel wall, its peripheral granules being permanently dispersed in the ectoplasmic fluid. Nevertheless, under such conditions, the cells perform amoeboid movements for several hours.

The independence of the surface movements from the state of viscosity of the endoplasm comes out more strikingly when isotonic or slightly hypotonic solutions (0.4 to 0.25 per cent) of citrate or oxalate are applied. The cells swell in inverse proportion to the concentration of the solution, this swelling being confined to the ectoplasmic layer which may eventually separate the endoplasmic capsule completely from the cell membrane. After a rapid succession of local solations and regelations, the plasmagel wall breaks down all over its surface and the granuloplasm becomes freely dispersed in the surrounding fluid (FIGURE 5). In spite of the total absence of a plasmagel layer, the cell membrane continues performing undulating movements for several hours. This would suggest (*1*) that membrane motility is less dependent on the presence of calcium in the external medium than is gelation of the endoplasm; and (*2*) that the surface movements cannot have been caused by contractions, or axial currents, of the endoplasm, but must be ascribed to autonomous changes of the cell membrane.

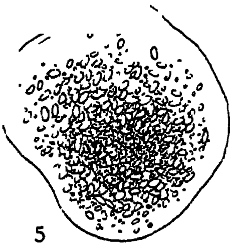
This total liquefaction and dispersion of the endoplasm is reversible when the cells are returned in time to a balanced salt solution containing calcium ions. Then the granules aggregate into a densely packed body and are surrounded again by a gelated wall which may be clearly distinguished from the fluid hyaloplasm. However, prolonged exposure to citrate or oxalate of hypo- or hypertonic concentrations eventually immobilizes and even dissolves the cell membrane. Disintegration of the membrane begins at one side of the cell and spreads slowly over adjacent regions, while the still intact portion of the membrane fails to contract, indicating that the decomposition of the surface film is preceded by a loss of its original elasticity.

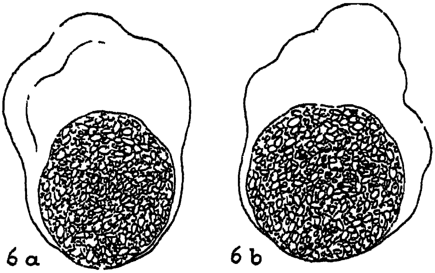
Similar observations have been made on cells which were exposed to alkaline or hypotonic media.

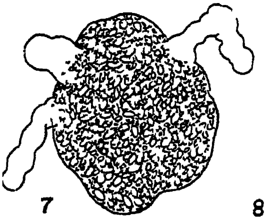
Cell Movements Despite Permanent Quiescence of the Endoplasmic Capsule. The reverse of the conditions in the preceding experiments, namely a permanent gelation of the endoplasmic wall, can be obtained by lowering the pH or increasing the relative calcium content of the culture fluid. Although these treatments simultaneously decrease the amount of ectoplasmic fluid, the amoeboid motility of the hyaline bulges or pseudopods is not suppressed.

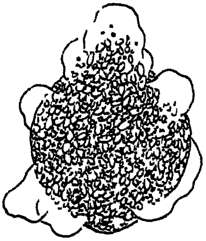
Another method of demonstrating the autonomy of the surface movements consists in subjecting the cells to standard solution containing alcohol (10 to 15 per cent). This treatment increases rather than decreases the volume of the hyaline layer, whereas the endoplasm contracts into a quiescent and sharply delineated ball which touches the cell membrane only in the posterior cell region (FIGURE 6). In the complete absence of endoplasmic solations, the cell membrane continues performing the characteristic undulating movements which shift the ectoplasmic fluid over the smooth surface of the endoplasmic capsule. It is mechanically inconceivable that any contractions of the capsule, or endoplasmic currents, might have caused the simultaneous appearance of more than one rotating bulge in the cell surface.

Attention may also be drawn to cells of the type of FIGURES 7–9 which


5




7


8

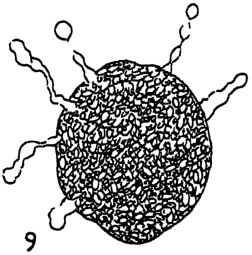

9

FIGURE 5. Complete solation and dispersal of the endoplasm under the influence of potassium oxalate, while the cell membrane continues to show amoeboid movements.

FIGURE 6. Exposure to 10% alcohol, preventing solation of the plasmagel but not the movement of the cell surface.

FIGURES 7–9. Pseudopod formation and fragmentation induced by brief exposure of the cells to alkali.

show the presence of numerous cylindrical pseudopods, devoid of endoplasm and moving independently of each other. The movements consist in bendings, longitudinal extensions and contractions, and peristaltic constrictions, traveling from the tip to the base of the hyaline pseudopods. These movements cannot be explained by a possible kinetic activity of the encapsuled endoplasm which remains quiescent. Such an outgrowth of the cell surface into numerous blunt pseudopods can be induced by a shock treatment with alkali.

Amoeboid Motility in Cell Fragments Lacking Endoplasmic Structures. If the conclusion is correct that it is not the reversible sol-gel process of the endoplasm but varying states of contraction of the cell membrane which cause cellular form changes and locomotion, then cell fragments lacking the endoplasm entirely should be motile as well. This is indeed the case.

An effective way of obtaining fragments of the desired composition consists in exposing cells of gastrula ectoderm briefly to alkali, which induces them to spread on the glass surface. After a while, the alkali effect wears off, the flattened cells contract and leave behind on the glass minute portions of their hyaline margin. Since the endoplasm is usually retracted into the bulky cell portion before the peripheral fragments become detached, the latter frequently contain no visible traces of structural cytoplasm. They may survive for many days and exhibit the various kinds of kinetic patterns observed in whole cells, *viz.*, spreading and locomotion of lamellar processes, alternate expansions and contractions of lobose or cylindrical pseudopods, and peristaltic constrictions passing over the surface of tubular fragments (Holtfreter, 1946a). With the aid of a surface of friction, hyaline fragments are capable of creeping forward like a vermiform cell. In cases where a fragment contains a few lipid granules, these move freely within the vesicle, rapidly dislocated by Brownian movement and by the irregular currents resulting from the undulations of the cell membrane. Whenever a portion of endoplasm is present, it forms, within the ectoplasmic fluid, a delineated body having a gelated surface layer and showing but very reduced translocations of the enclosed granules. There occur all intermediary forms between completely hyaline vesicles and others which are almost entirely filled with endoplasmic material. The former are usually much more motile than the latter.

When exposed to increasing concentrations of standard solution, the hyaline vesicles shrink, their motility slows down and eventually stops. This process may be reversed. In still higher concentrations of this solution, the vesicles die and become filled with basophilic granules, showing unrestricted Brownian movement. They can be further dehydrated by various agents such as acids, or concentrated solutions of calcium chloride, basic dyes, or heavy metal salts, all of which transform the vesicle into a thin, granulated film firmly attached to the glass surface (FIGURE 10).

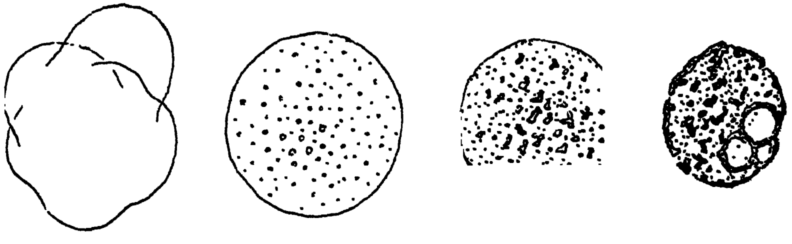


FIGURE 10. Exposure of a hyaline cell fragment to 10^{-1} M $CaCl_2$, causing shrinkage, granulation, and collapse of the vesicle in the form of an adhesive basophilic film.

From these observations, it may be concluded that the different modes of amoeboid form changes can be executed by cell portions which lack entirely the inner protoplasmic material capable of a reversible gelation. It appears that this conclusion applies also to *Amoeba*, for Chambers (1924) observed that cut-off pseudopods of *Amoeba* were capable of ingesting food and moving about in the typical amoeboid manner, although they were entirely free from the visible granules of the parent body. Taken together with the evidence of the preceding paragraphs, these data indicate that the endoplasmic gel-sol formation occurring sometimes in whole cells is merely superposed upon the kinetic activity of the cell membrane. The present material furnishes no evidence that the plasmagel capsule has a contractility of its own. Whenever the endoplasmic material changed from a spherical into a cylindrical shape, showing peristalsis, it was tightly enclosed by the cell membrane, suggesting that it was the well-attested contractile power of the latter which had molded the endoplasm. If these two structures were separated by a cushion of ectoplasmic fluid, the outer membrane continued performing amoeboid movements, whereas the surface of the endoplasm remained immobile.

The periodical solations of the plasmagel apparently result from chemical reactions between the endoplasm and the supernatant ectoplasmic fluid. This process seems to involve the exchange of electrolytes, with calcium ions playing a predominant role in controlling the viscosity of the structural endoplasm. While in some instances liquefaction and dispersion of the granulated endoplasm proceed slowly, comparable to the melting of a submerged sheet of ice containing particles, in other cases the plasmasol may rush rapidly through the opening in the plasmagel wall. This acceleration of flow, in contrast to the liquefaction proper, appears to be caused by the contraction of an external covering. Since forceful endoplasmic eruptions occurred only if the posterior portion of the endoplasm was closely associated with the undulating cell membrane, the contractile force driving the plasmasol toward the hyaline cap is probably furnished by the outer membrane.

Constriction of the Cell Membrane Resulting in Cytoplasmic Division. Isolated embryonic amphibian cells may exhibit three kinds of division phenomena: (*a*) mitotic nuclear division associated with cytoplasmic fission; (*b*) division of the nucleus into two or more nuclei without segmentation of the cell body; (*c*) cellular fragmentation in the absence of nuclear division. The latter phenomenon may occur within a normal embryo, for instance in the flask-shaped blastoporal cells, after they have reached their final destination in the anterior region of the archenteron. The formation of blood platelets has been attributed to the budding-off of cytoplasmic fragments from large mother cells. Lewis (1942) ascribes cytoplasmic division to the same mechanism which produces the constriction waves in migrating cells. The following data support this concept, although they do not agree with the idea of Lewis that it is the plasmagel which performs the constrictions.

In the cylindrical cells mentioned above, it was frequently observed that a constriction wave became stationary, that the constricting ring cut progressively deeper into the cell body and finally caused a fragmentation of the cell. The products of division were usually very unequal in size and composition. The nucleus remained inert and intact and came to lie in one of the daughter bodies, both of which could survive for several days. That it was again the cell membrane which provided the mechanism of constriction was indicated by observations on the cell fragments mentioned above. Very small hyaline fragments containing neither chromatin nor any other microscopically visible internal structures, would divide spontaneously into two bodies which continued performing amoeboid movements.

Cytoplasmic division in the absence of a nucleus has been observed in the eggs of the axolotl, the second spindle of which had been removed by means of a micro-pipette (Jollos and Peterfi, 1923). Fankhauser (1934) observed in polysperm merogons of *Triton* that cell walls may form around astrospheres lacking chromosomes and that segmentation of the cytoplasm may occur even independently of any nuclear or astral activity. On the other hand, division of the sperm nuclei or of the accessory astropheres was not necessarily associated with a corresponding cleavage of the cytoplasm. In centrifuged eggs of the sea urchin which had been activated by hypertonic sea water, Harvey (1938) found that cleavage could take place while the egg nucleus was still intact, which suggests that "the cleavage of an egg and the nuclear changes usually accompanying it are quite separate phenomena." Non-nucleate halves of the eggs of sea urchins and of *Chaetopterus*, which had been treated with a parthenogenetic agent, performed amoeboid movements and then fragmented into cells of irregular sizes containing astrospheres but no chromatin (Harvey, 1936, 1938, 1939). In non-nucleate pieces of the starfish egg which had been subjected to parthenogenetic agents, Chambers (1924) observed the budding-off of fragments, some of which possessed no cytasters. There is no observational evidence that the cytoplasmic fissions occurring in bacteria and some algae lacking a distinct nucleus are associated with the presence of cytasters. These observations make it improbable that cytaster formation is a necessary prerequisite for cytoplasmic division.

It would appear that regional differences of composition of the protoplasm, associated with the presence and division of astrospheres, tend to localize and arrest the autonomous constriction movements of the cell surface in an equatorial zone. Here, the constriction would proceed to form a cleavage furrow, perhaps because the underlying cytoplasm is comparatively more liquid and, hence, less resistant than elsewhere, or because some chemical or electric properties of this region induce a locally stronger contraction of the cell membrane. Chalkley (1935) noticed in dividing amoebae that the plasmagel layer in the region of the deepening furrow becomes thinned out, liquefies, and streams away in oppositely directed axial currents. Erlanger (1897), Spek (1918), and other investigators have recorded the streaming of superficial liquid cytoplasm from the poles toward the plane of cleavage. In centrifuged sea urchin eggs which were subsequently fertilized, the first cleavage furrow sank in more rapidly on the hyaline than on the densely granulated side of the egg (Chambers, 1924).

These observations do not support the view that the viscous plasmagel layer is actively engaged in the cleavage process. According to observations of Motomura (1935) on sea urchin eggs, the external plasma membrane covering the furrow cuts through the viscous cortical layer without dislodging it from the egg surface. Dan, Yamagita, and Sugiyama (1937), from observations on the displacement of kaolin particles attached to the plasma membrane of sea urchin eggs, concluded that the surface area over the spindle poles expands throughout the cleavage process while the prospective furrow region at first shrinks, then increases markedly when it is drawn into the depth of the furrow. Similar conditions apply to the amphibian egg (Schechtman, 1937). However, the opaqueness of the latter egg makes it impossible to watch the behavior of the cytoplasm. That the ingression of the furrow results from the constriction of a densified region of the cell membrane seems to be suggested by the fact that the floor of the furrow is more resistant to chemical and mechanical injuries than is the cell surface over the spindle poles (Just, 1922: Chambers, 1938). This would agree with the conditions in hyaline pseudopodia and cell fragments where it can be noticed that the constricting rings acquire temporarily a higher refractivity than the expanding areas, suggesting alternate states of density of the cell membrane. In dividing fibroblasts, the opposite poles of the cell are known to bulge out into rapidly moving blebs which recall the hyaline blisters formed under the influence of alkali, mechanical irritation, or other agents which weaken and liquefy the cell membrane.

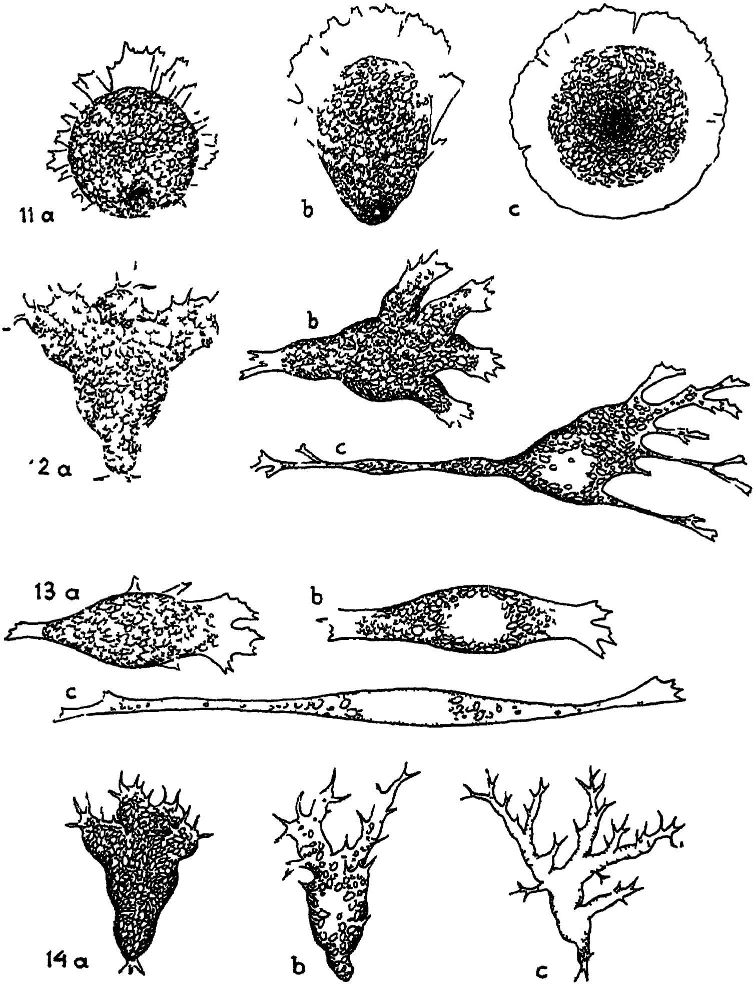
The above considerations make it understandable that external stimuli, causing a local reduction of density and contractility of the cell membrane, will at first increase the permeability of this layer, then induce the formation of pseudopods and hyaline blisters, and finally cause a pinching-off of the out-bulging portions of the cell. This phase may be followed by cytolysis. Most parthenogenetic agents produce cytolysis when applied for longer periods or at higher concentrations. Comparable with the artificially induced formation of hyaline protuberances and their detachment in immature sea urchin eggs (Runnström, 1928), a pinching-off of amoeboid vesicles free of granulated endoplasm may occur in gastrula cells and in early erythrocytes of amphibians which have been briefly exposed to alkali or other agents which at higher concentrations decompose the cell membrane. Mechanical irritation may likewise cause a fragmentation of the embryonic cell. The main difference between the constrictions associated with amoeboid movement and those leading to cell division, or fragmentation, appears to be that the former are reversible and successively involve various regions of the cell surface, whereas the latter remain localized and cut progressively deeper into the cytoplasm.

Function of the Cell Membrane in the Process of Cellular Differentiation. By way of controlling cellular locomotion and cytoplasmic division, the cell membrane plays an essential role in bringing about the morphogenesis of the different types of cells. A differentiating cell acquires its new shape through the elaboration of new patterns of amoeboid activity, a phenomenon usually accompanied by the development of regional, cell-specific differences of adhesiveness of the surface film. In order to materialize these properties effectively, most cells need an external surface of contact. A few examples may illustrate the situation. The cells to be discussed had been derived from urodele embryos and were cultured singly in physiological salt solution (Holtfreter, 1946b, 1947b, c).

Whereas the different cell types from pre-neurula stages are, after isolation, more or less spherical, showing only slight tendencies of spreading, of becoming cylindrical, or forming slender pseudopodia, their morphogenetic behavior becomes more distinctly cell-specific after they have passed these stages. Isolated prospective epidermis cells from a neurula tend to flatten shield-like against glass or living tissues (FIGURE 11), while the individual cells from the neural plate tend to stretch themselves into long vermiform bodies (FIGURE 4). On the other hand, neuroblasts of the type of Rohon-Beard's cells, as well as the cells from the neural crest, shortly after their isolation from early tail bud stages, adopt a shape somewhat resembling that of an *Actinia* (FIGURES 12a and 14a). In the absence of a supporting substratum, their plump anterior portion may project hyaline tapering pseudopods capable of elongating, contracting, and bending around in curves. The posterior cell portion is conical and can change its shape by way of constrictions and longitudinal stretchings and contractions. These cell-specific differences of kinetic behavior are a reflection of the fact that from now on the various cells are determined cytologically.

Evidently, the different cell shapes arise from the activity of the cell membrane, not the endoplasm. The motile filopods of an early neuroblast contain merely ectoplasmic fluid; the anterior portion of the vermiform neural plate cell may exhibit cylindrical elongations, bendings, and constrictions while not underlain by the endoplasm; the epidermis cell spreads and progresses over the substratum by way of extending a hyaline margin which is distinctly set off from the central body of granulated endoplasm.

With the substitution of lamellar or filiform pseudopods for the earlier rounded lobopods, the amount of ectoplasmic fluid is generally reduced. At the same time, the reversible solation of the plasmagel becomes less frequent and is confined to the increasingly smaller areas which directly border at a hyaline protuberance, while the bulk of the



FIGURES 11–14 Successive stages of differentiation in an isolated epidermis cell (FIGURE 11), in a neuroblast (FIGURE 12), a myoblast (FIGURE 13), and a mesenchyme cell (FIGURE 14)

endoplasm remains unaffected. Local outbreaks of a peripheral portion of the granuloplasm have been observed in the margin of flattened epidermis cells and in the tongue-shaped advancing pseudopods of neuroblasts, myoblasts, and mesenchyme cells. The process is readily recognized because of the sudden acceleration of movement which the granules undergo when pouring into the ectoplasmic fluid. It should be emphasized, however, that, as in the case of lobose or tubular pseudopoda which project freely into the external medium, the flattened pseudopods of attached cells move just as well in the absence as in the presence of gel-sol formations.

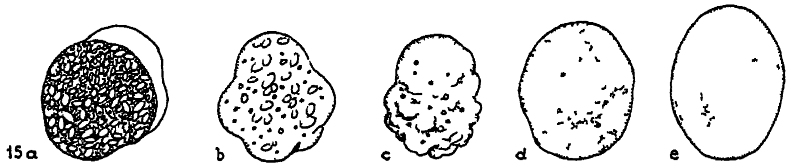
The movements of the cell membrane in flattened cells are less coordinated and on a smaller scale than are those in non-attached cells, where they pass in large waves over the whole cell body. Peristaltic constrictions do occur in differentiating neuroblasts, myoblasts, leucocytes, and mesectoderm cells, but they become less conspicuous the more the whole cell flattens against the substratum. Instead, the hyaline margin of an attached cell is engaged in locally independent and irregularly alternating movements of expansion and contraction which produce a constant change in the outlines of the serrated periphery. Furthermore, ring patterns and ruffles may move over the surface of the hyaline margin. Such ruffles are well-known features of the laminar processes in monocytes and macrophages of vertebrates. Fauré-Fremiet (1929, 1930), who observed them in the freely extended membranous pseudopods of choanoleucocytes, stresses the fact that the movements proceed while the centrally located endoplasm remains quiescent. He therefore rejects the idea that the movements of the hyaline processes are caused by variations of the gel-sol ratio of the endoplasm. This conclusion is supported by our observation that the same kind of movements may occur in isolated fragments of the hyaline margin which contain no endoplasm at all. In migrating cells, the endoplasm merely follows the movement of the advancing margin, by filling up successively the hyaline space, thus securing the ground gained by the activity of the ectoplasmic cell portion. If the total sum of periodical extensions outweighs that of contractions, the flattened pseudopod will continue moving forward in one direction.

All cell types mentioned above possess an antero-posterior polarity. The relatively inert posterior pole corresponds to that side which in the blastula stage was facing the outside. In epidermis cells which are merely spreading, the posterior pole takes up an apical position (FIGURE 11c). In migrating cells, the posterior portion is dragged behind and contracts periodically, frequently forming a tail knob, while the advancing anterior region may remain more or less unipolar (myoblasts) or may branch out into several independently moving pseudopods (neuroblasts, mesectoderm cells, leucocytes). If the posterior pole becomes firmly attached to a substratum and the anterior cell portion continues advancing, the whole cell may become stretched into a long ribbon (myoblasts, FIGURE 13c), or only a tail portion is spun out into a fibrous process (neuroblasts, FIGURE 12c). The transformation of the dendritic anterior protuberances of a neuroblast into long fibrous processes is only possible if the bulk of the cell remains sessile and the pseudopods continue advancing by means of their undulating hyaline end-plate. Leucocytes exhibit a pronounced mobility because of lack of adhesiveness of their posterior region and a great agility of their anterior region.

It appears that, with progressive differentiation, the cell membrane becomes more firmly attached to the underlying plasmagel and that this reduces or inhibits its motility. This condition is especially conspicuous in isolated epidermis cells which have been cultured for several days.

Their entire endoplasm forms a gelated layer closely associated with the surface film, which becomes immobile. In fully extended spindle-shaped myoblasts and in dendritic mesenchyme cells, the major part of the cell surface becomes progressively quiescent, motility being confined to the tips of the pseudopods which have remained free of endoplasm.

Congelation of the whole cell occurs in a pronounced form in erythrocytes (Holtfreter, 1947c). Like all other cells of the amphibian embryo, the erythroblasts represent at first globular cells provided with rotating lobopods which may undergo local gel-sol formation (FIGURE 15a). When the hemoglobin becomes visible and the yolk is almost absorbed, the ectoplasmic fluid slowly disappears and the cell membrane settles down over the entire surface of the plasmagel capsule, without, however, being arrested in its kneading movements (FIGURE 15b). Unlike the spreading movement in epidermis cells, the flattening of the erythroblast into a discoidal body occurs in the absence of any substratum and is executed simultaneously by both the cell membrane and the endoplasm. After the cell has flattened, it remains immobile, and Brownian agitation of the endoplasmic inclusions becomes less discernible. However, the process of immobilization can be reversed, if the early erythrocyte be exposed to liquefying agents, such as salt solutions hypotonic or of high pH (FIGURE 16a–e). Under such conditions, the rim of the discoidal body develops amoeboid protuberances, the cell rounds up and exhibits the kneading movements of its earlier stage, and finally the outer membrane is lifted off from the plasmagel layer and forms once again undulating lobopods filled with liquid cytoplasm now containing hemoglobin. The lobopodia may become pinched off at their base and form mobile fragments (FIGURE 16d). With extended treatment the plasmagel layer becomes liquefied as well, at first locally and reversibly, then all over its surface. This causes the endoplasmic inclusions to be freely dispersed within the now spherically swollen cell. At this stage, the cell membrane becomes immobile and so porous that hemoglobin and even larger particles includ-



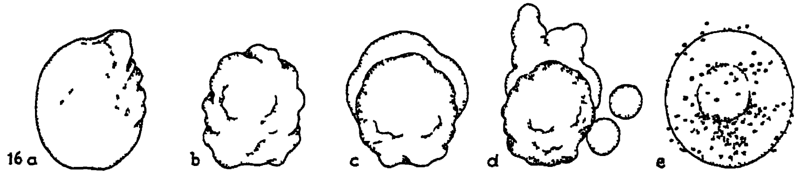


FIGURE 15. Successive stages of differentiation of an erythroblast

FIGURE 16 Retrogressive form changes of an erythrocyte following exposure to alkali.

ing the nucleus can penetrate it. The final stage of cytolysis is accompanied by the appearance of precipitation granules in the cytoplasm.

A similar sequence of reactions to liquefying solutions can be obtained with other differentiated cells, *e.g.*, leucocytes or mesenchyme cells. The treated cells will return to their earlier mode of motility by replacing the filiform or dendritic processes by large undulating lobopods and they eventually die from being flooded with the immersion fluid.

Thus, cellular differentiation which is associated with partial immobilization and a restriction of amoeboid activity to localized pseudopods, seems to involve a certain amount of dehydration. During the period of consolidation of the final cell shape, the inner cytoplasm may find the proper conditions to develop structures of increasing stability, as, for instance, oriented fibrils. Observational data and concepts as regards the possible genesis of cytoplasmic structures may be found in the publications of Schmidt (1937), Frey-Wyssling (1938), Picken (1940), Schmitt, Hall, and Jakus (1943), and Lawrence *et al.* (1944). If the above considerations be accepted, the appearance of an intracellular cytoskeleton is the consequence rather than the cause of the external form changes in differentiating cells.

Observations on the Composition, Structure, and Physiological Reactions of the Cell Membrane

In order to understand the various functions of the cell membrane, it is necessary to inquire into its chemical composition and physical properties. The pertinent data, although obtained with different methods and on different material, appear to conform sufficiently to be pieced together into a comprehensive, though somewhat sketchy concept. (See the reviews by Harvey and Danielli, 1938; Danielli, 1942; Schmitt, 1941, 1944; Schmitt and Bear, 1939; Monné, 1946.)

Recent investigations afford ample support for the idea, advanced by early workers on cellular permeability, that the cell is bounded by a film containing lipids and probably also proteins. Not only the surface film of the different somatic cells so far studied, but also the outer protoplasmic layer of eggs appear to consist principally of an organized lipoprotein structure.

Runnström and collaborators (1928, 1943, 1945) have shown that, when unfertilized sea urchin eggs are exposed to lytic agents, such as alkali, fat solvents, high temperatures, hypotonic sea water, or merthiolate, their cortical layer bulges out into hyaline blisters and pseudopods which may become detached in the form of "lipoid vesicles." These formations resemble the pseudopods and hyaline fragments in amphibian cells, which arise either spontaneously or under the influence of various cytolytic agents. The cortical layer of the living normal sea urchin egg is positively birefringent in a radial direction, suggesting an ultrastructure comparable to that in myelin formations, where the elongate phosphatide molecules are assumed to form concentrically arranged bimolecular leaflets, the long axes of the molecules paralleling each other and lying perpendicularly to the surface of the body (Runnström, Monné, and Broman, 1943; Monroy, 1946, 1947). Through the action of fat solvents, the birefringence of the cortex may either disappear completely or reverse its sign, becoming a radially negative one. These observations suggest that the cortex contains alternate lamellae of lipid and protein molecules, the latter being extended in a tangential direction and attached laterally to the polar groups of the lipid leaflets (Monné, 1946). It has not been possible to decide whether the entire birefringent cortex, which is quite thick (about 1μ), or only its peripheral portion should be identified with the selectively permeable cell membrane.

Under dark-field illumination, the birefringent cortex of unfertilized sea urchin eggs has a yellow-orange color, changing into white and bluish-grey when this layer is expanding into the pseudopods mentioned above (Runnström, 1928). Monné (1941), who studied the color variations in the protoplasmic constituents of a great variety of cells, concludes that they reflect alterations in the state of hydration of the colloids involved. According to Monroy and Monroy Oddo (1946), the change and final disappearance of color in the cortex of the sea urchin egg is associated with a progressive reduction of the positive birefringence of this layer, indicating a dispersal or dissociation of the constituent lipid molecules, probably because of an increased solvation. It is interesting to note that a disappearance of color and birefringence can be produced not only by fat solvents, hypotonicity and other lytic agents, but also by a slight compression of the egg (Monroy and Monroy Oddo, 1946). Upon return of the egg to normal conditions, the positive birefringence may reappear, although the cortex of compressed eggs fails to recover completely, exhibiting patches of disarrangement of its molecular constituents.

On the basis of these observations, one may assume that the hyaline blisters which can be produced by both chemical and mechanical agents, are the result of a locally increased solvation and porosity of the cell membrane, and the accumulation of water beneath it. In a similar way, the spontaneous local expansions occurring in the cell membrane of amoeboid cells appear to involve a higher state of solvation as compared with that of contracting regions. In a hyaline pseudopod, the out-bulging portions are less refringent than are the constricted ones. Hyman (1917) found that the tip of a pseudopod in *Amoeba* is more susceptible to injurious chemicals than is the rest of the cell surface.

A cortical fine structure comparable to that in the echinoderm egg has been reported to occur in various other cells. Combined chemical, leptoscopic, and polarization-optical studies on the ghosts of hemolyzed erythrocytes suggest that their limiting membrane consists of alternating lamellae of lipids and proteins, their arrangement resembling the molecular structure of the myelin sheath of nerve fibers (Schmitt, Bear, and Ponder, 1936, 1938; Waugh and Schmitt, 1940). A corresponding lipoprotein structure has been assumed to exist in the cell membrane of living embryonic chicken cells (Hobson, 1941), of differentiated ganglion cells of the frog (Chinn, 1938), and of living erythrocytes and spermatocytes of invertebrates (Monné, 1941, 1946). According to Fauré-Fremiet (1929), the hyaline pseudopods of choanoleucocytes are birefringent when in a dehydrated and agglutinated condition. Alcohol and ether, in contrast to acetone, destroy this property, indicating that it is due to the presence of phospholipids. I observed birefringence in the surface layer of living amphibian eggs and in the cell membrane of the different isolated cells and hyaline fragments discussed above, but I have not yet succeeded in obtaining more exact data on this phenomenon.

The concept of an organized lipoprotein structure of the cell membrane is supported by a variety of experimental results concerning the permeability, surface tension, solubility, and electric properties of the cell surface (see Harvey and Danielli, 1938). The concept agrees, furthermore, with the findings on the chemical composition of this structure. Monroy and Monroy Oddo (1946), on the basis of quantitative polarization-optical studies, conclude that the birefringent cortex of the echinoderm egg consists predominantly, if not entirely, of phosphatides and cholesterol. In the red blood cell, practically all the lipids are concentrated in the outer membrane, which constitutes the posthemolytic residue (Erickson *et al.*, 1938). According to the analyses of Parpart and Dziemian (1940), this membrane contains lipids and proteins at a ratio of about 1:1.7, the lipids present consisting almost entirely (82 to 98 per cent) of phosphatides, especially cephalin, and of cholesterol.

The presence of phosphatides in the cell membrane of embryonic amphibian cells is suggested by the observation that this structure is readily destroyed by cobra venom. Increasing concentrations of alcohol render the membrane highly permeable and may eventually dissipate it entirely. Although the cell membrane is less susceptible to the absence of calcium ions in the immersion fluid than is the coat of the amphibian egg, its amoeboid activity and stability seem to depend on calcium bound to the lipoprotein components (see p. 711). Monné (1946) draws attention to the fact that the surface of many cells has been found to be rich in calcium and magnesium and that nucleic acids are also present. He suggests that the phosphoric acid groups of the phosphatides and of the nucleic acids are held together by calcium molecules.

Comparison between the Behavior of Lipid Models and the Cell Membrane

At this point, a comparison between the behavior of the living cell membrane and that of artificial films of lipids and lipid mixtures may promote the further analysis of our problem.

Fauré-Fremiet (1925, 1929, 1930) and Runnström (1928) observed certain similarities between the kinetic, chemical, and polarization-optical properties of the cell surface and those of myelin tubes from lecithin or cephalin. Having found that the hyaline pseudopods of amoebocytes are anisotropic and contain large proportions of phosphatides, Fauré-Fremiet (1929, 1930) suggested that their movements could be interpreted as resulting from localized variations in the degree of hydration of surface layers of oriented phosphatide molecules. The extent of hydration and, consequently, the expansions and contractions of the pseudopods were assumed to be controlled by pH and by the kind and proportion of the cations present.

Some gross similarities between the appearances and kinetics of living cells and those of models of lipids and their mixtures with other substances have been noticed by several other workers (Herrera, 1932; Crile, Telkes, and Rowland, 1932; van Herwerden, 1933). It was, however, the aspect of molecular structure and permeability rather than that of motility of the cell surface which led to the more intensive studies on lipid models. Harvey (1912) produced vesicular bodies of lecithin and proteins which reacted to neutral red, saponin, and mechanical pressure not unlike sea urchin eggs. Danielli (1936, 1945) and Harvey and Danielli (1936) studied artificial lipoprotein films from the viewpoint of elasticity, interfacial tension, and adsorption power. In a series of important contributions, Bungenberg de Jong and co-workers (1932, 1935, 1937) investigated the effects of physical agents, of hydration, pH, electrolytes, proteins, and various other substances on the behavior of lipid models. The concepts on coacervation which emerged from these studies represent a very instructive guide for an interpretation of the structural and functional properties of the cell membrane. No less fruitful for the cytologist has become the extensive work on monomolecular layers, which is connected with the names of Adam, Langmuir, Rideal, and Schulman.

The writer has become interested in lipid models because their reactions to various chemical and physical agents seemed to explain some processes connected with Golgi bodies and cellular vacuoles (Holtfreter, 1946c). At the same time, these experiments disclosed phenomena which appear to promote an understanding of the cell membrane. This dual applicability of the models is not surprising since the structure and functions of the cell membrane are supposed to resemble in many ways those of the lipoprotein film surrounding intracellular vacuoles (Scarth, 1927, 1940; Bungenberg de Jong, 1932, 1935; Frey-Wyssling, 1938). Some of these observations which have a bearing upon our subject matter will now be discussed. The morphological and kinetic behavior of living cells and hyaline cell fragments was compared with that of myelin bodies from crude lecithin, which were exposed simultaneously to immersion fluids of different compositions. The lecithin used was in an advanced state of rancidity and contained admixtures of cephalin, free fatty acids, and probably other lipids. This impurity of the "lecithin" preparation was not considered disadvantageous, since it appears that the biological membranous structures likewise represent mixtures of different fatty substances.

When immersed in distilled water or dilute Ringer, the lecithin developed the characteristic myelin tubes and vesicles which, with further uptake of water, grew in length and diameter. The thickness of the tubular walls and the extent of elongation varied with the initial mass entering into the swelling process. The larger bodies usually gave rise to complex systems consisting of vesicles which included other vesicles and tubes of various size. Small grains of lecithin rapidly became spherical and developed a large cavity filled with an optically homogeneous fluid, while the external wall was thinned out into a film having about the diameter of the cell membrane. In their optical appearance (ordinary and polarized light), these small vesicles were practically indistinguishable from living contiguous hyaline cell fragments.

Myelin vesicles, stretched by glass needles into a long tube and released, either contracted or broke down into several portions which rounded up into spherical fragments. When embedded in olive oil, hydrated myelin formations could be drawn out into very thin fibrillar and anastomosing structures resembling the processes of mesenchyme cells. The surface film of a myelin vesicle was capable of sealing up a hole made by a glass needle, and of restoring its continuity when penetrated by vacuoles expelled from the interior.

Reactions of Myelin Bodies to Variations of pH. In analogy with the behavior of hyaline pseudopods, myelin bodies form tubular shapes only within a medium pH range, acidulation causing shrinkage and granulation, alkalinization, a vesicular swelling, fragmentation, and dispersal of the bodies. FIGURE 17 shows diagrammatically the relationships between the pH of the immersion fluid and the type of myelin formations which developed at the periphery of a lump of lecithin within the period of an hour. The immersion fluid consisted of distilled water to which HCl or KOH respectively had been added. It will be noticed that the swelling of the periphery increases in direct proportion to the raising of pH. At the lowest levels (pH 2.0–2.2), only a thin surface layer swells into a homogeneous transparent substance. At higher levels, up to about pH 4.0, myelin figures begin to appear, but shrivel up into granules when with further elongation they come into direct contact with the acid medium. Granulation is no longer detectable above pH 4.2. Instead, swelling continues and increases gradually with rising pH. Within the pH range of about 5.0 to 9.0, there are no marked differences in the size and shape of the strongly expanded myelin bodies. At levels above pH 9.4, the tubes become increasingly spherical and may form numerous thin-walled bulges which are pinched off as separate vesicles or may burst and disintegrate (FIGURE 18).

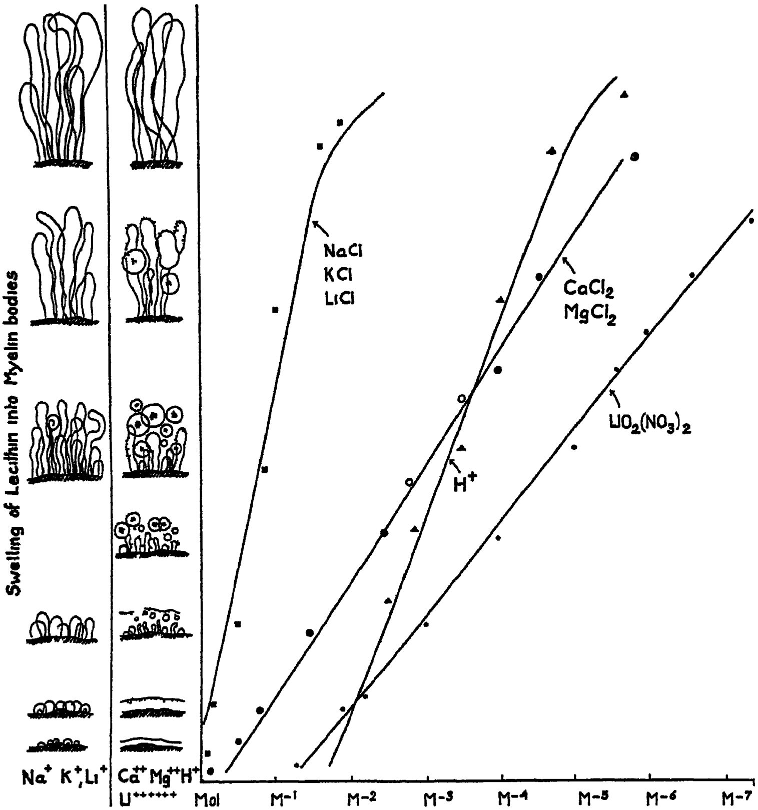
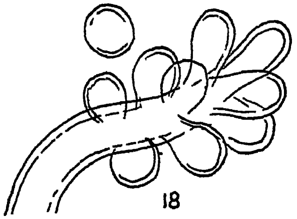
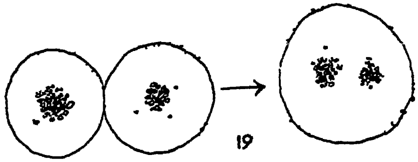


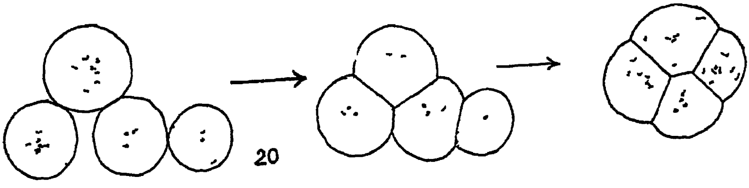
FIGURE 17 Effects of different concentrations of hydrogen ions and various metal ions on the swelling of crude lecithin The effects have been diagrammatically indicated by the relative thickness of the marginal zone made up of myelin bodies

When a fully expanded myelin vesicle is transferred from a neutral into an acid medium, it shrinks and the internal fluid gives rise to granules showing Brownian movement. With increasing acidity, the granules become coarser and the vesicle shrivels up and collapses to form a thin, granulated film firmly attached to any contact surface (FIGURE 20a). Adhesiveness of the lecithin vesicles to each other or to glass becomes apparent at about the same pH level (∓4.2) when granules are being formed The extent of flattening of the bodies over a substratum increases with the rate of dehydration.

In an analogous manner, the size, shape and adhesiveness of the hyaline processes in embryonic cells depend upon the pH of the immersion fluid. In standard solution which is rendered alkaline, the cell swells and its filopodia are transformed into large bulges which may become detached as vesicles and eventually disintegrate. Cellular aggregates break up into single cells at pH above 9 4 and below 4.2. Beyond this range,


18


19


20

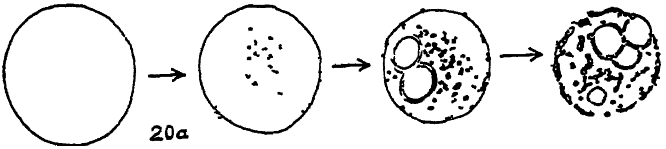

20a

Figure 18 Accessory buds and fragmentation occurring in a myelin tube under the influence of alkali

Figure 19 Coalescence of two myelin vesicles in a 5×10^{-4}M solution of $CaCl_2$

Figure 20 Aggregation of myelin vesicles in a 10^{-2}M solution of $CaCl_2$

Figure 20a Shrinkage, granulation and collapse of a myelin vesicle under the influence of acid or other desolvating reagents

amoeboid motility ceases, and the cells eventually die. On the acid side of this range, the dead cell membrane becomes increasingly adhesive to glass or to lecithin vesicles, and the content of hyaline cell fragments precipitates into granules. Strong acids cause complete shrinkage and collapse of the fragment into a coarsely granulated film spreading on glass (FIGURE 10)

Effects of Metallic Ions. The rate of swelling of lecithin in solutions of the chlorides of Na, K, Li, Ca, and Mg varies with the salt concentration and with the valence of the cations, bivalent cations having a much stronger inhibiting effect than monovalent cations (FIGURE 17).

In molar concentrations, all chlorides completely suppressed myelin formation. Equimolar solutions of NaCl, KCl, and LiCl apparently had an identical inhibitory effect which was no longer noticeable at concentrations below 10^{-2}M, while that of $CaCl_2$ and $MgCl_2$ could still be traced at concentrations as low as 10^{-5}M. Morphologically, the effect of the various metal ions was similar to that of hydrogen ions. In higher concentrations of the bivalent cations, only small spherical myelin bodies were formed. Length and width of the bodies increased with a decrease of the salt concentration. When expanded myelin bodies were transferred from distilled water to increasingly concentrated solutions of any of the chlorides they shrank, developed vacuolar and granular structures both within their internal fluid and in the outer membrane, and finally shrivelled up into compact bodies or, if attached to a substratum, into a reticulo-granular film (FIGURE 20a).

In accordance with its strongly dehydrating action, calcium chloride rendered myelin bodies adhesive at concentrations as low as 5×10^{-5}M. When, in more concentrated solutions of $CaCl_2$ or $MgCl_2$ (3×10^{-4}M and stronger), two or more lecithin vesicles were brought into contact with each other, they either flattened one against another, forming close-packed aggregations comparable with those of living cells (FIGURE 20) or fused together into a single sphere with a complete disappearance of the separating walls (FIGURE 19). Such fusions could also be obtained with the application of acid.

In a comparable way, the rate of water uptake of living amphibian cells, or their hyaline fragments, varies markedly according to whether the culture fluid contains mono- or bivalent cations. At concentrations which cause complete dehydration of myelin bodies, all the chlorides used transform a hyaline fragment into an adhesive film studded with granules and having the appearance and basophilic properties of a collapsed lecithin vesicle. The embryonic cells attain maximal lobopod formation and amoeboid motility in relatively strong solutions of NaCl or LiCl (around 5×10^{-2}M). Lower salt concentrations cause cytolysis preceded by an extreme expansion, loss of elasticity and partial dispersion of the cell membrane. In pure solutions of $CaCl_2$ or $MgCl_2$, of any concentration, the cells cannot be kept alive for a longer period. At concentrations around 5×10^{-2}M, these salts produce permanent gelation of the endoplasmic wall, slowly moving pseudopods, and finally total coagulation of the cell. Like myelin vesicles, living embryonic cells are non-adhesive in standard solution lacking calcium. The minimal amount of $CaCl_2$ necessary for cellular aggregation in a 0.35 per cent NaCl solution at pH 8.0 is about 10^{-5} to 10^{-6}M.

A process comparable to the coalescence of adjacent lecithin vesicles has been observed in blastula cells of *Amblystoma*. When exposed to standard solution of double strength or having a highly increased ratio of $CaCl_2$, the separating walls of a number of cells disappeared and multinuclear syncytia were formed. A mutual or multiple coalescence may occur both in normal and pathological cells (fusion of gametes; syntrophoblast of the placenta; giant cell formation in cultured tumor cells, etc.). The phenomenon seems to involve the disappearance of a layer of bound water at the interface of the cells in contact.

Several other metal compounds, such as gold chloride, mercuric chloride, silver nitrate or osmium tetroxide likewise strongly dehydrate both myelin bodies and hyaline cell fragments, while sodium oxalate and citrate influence myelinitic growth no more drastically than do corresponding concentrations of sodium chloride. It seems that the dehydrating power increases with the valence of the cation. Experiments carried out with uranyl nitrate showed this compound to be still more effective than the bivalent cations, since its inhibitory action upon myelinitic swelling could be noticed at concentrations as low as 10^{-7}M (FIGURE 17).

Lecithin vesicles which have become completely dehydrated through the action of heavy metal ions can no longer be reversibly hydrated and are more or less insoluble in alcohol and ether. This suggests that the well-known toxic and cytolytic action of heavy metals may arise, in part at least, from their irreversible combination with the phosphatides of the cell. According to Waugh and Schmitt (1940), copper salts at a 10^{-5}M concentration render the envelopes of erythrocytes largely insoluble in organic solvents and insensitive to pH and to the absence of electrolytes. It is thought that the heavy metal ions form cross-linkages with the lipids and proteins of the cell membrane, thereby disarranging the molecular pattern and increasing the permeability of this structure (see, also, Dawson and Danielli, 1938).

Palmer and Schmitt (1941) have analyzed the x-ray diffraction patterns of cephalin emulsions which had been subjected to the action of various electrolytes. $CaCl_2$ was found to be much more effective in reducing the long-period spacings of the lipid structure than were equivalent concentrations of NaCl or KCl. The pronounced desolvating effect of calcium ions was ascribed to their combination with the phosphoric acid groups of the bimolecular lipid leaflets, and the ensuing expulsion of water from between the polar interfaces of the leaflets. A similar mechanism was assumed to account for the desolvating and flocculating action of basic proteins, such as histone, upon phosphatide emulsions (Palmer, Schmitt, and Chargaff, 1941). This concept appears to give a valuable clue to the interpretation of cellular adhesiveness. Schmitt (1941) observed that when small amounts of histone or thorium salts were added to a suspension of red blood cells, the cells were tightly drawn together and formed aggregates. The phenomenon was explained as resulting from an attachment of the introduced cations to the acidic groups of the cell membrane, causing a desolvation and linking-up of the molecules in the adjacent cell surfaces.

Antagonistic Effects of Monovalent and Bivalent Cations. Prolonged exposure of lecithin to molar concentrations of $CaCl_2$ or $MgCl_2$ renders the substance incapable of swelling again in water. However, myelin vesicles which have suffered only a partial dehydration in less concentrated solutions of these salts will swell again when the solutions are diluted with water. The desolvating effect of the bivalent cations can likewise be counteracted by monovalent cations. Thus, moderately dehydrated myelin bodies which have remained unchanged in a 10^{-4}M solution of $CaCl_2$ for several hours will grow out into long myelin tubes when any of the chlorides of Na, K, or Li are *added* to the solution, provided the total salt concentration does not rise above 10^{-1}M. This indicates that the combination of the bivalent cations with the phosphatide molecules of the external membrane is, to a certain extent, reversible, and that the bivalent cations can be expelled and replaced by monovalent cations whose desolvating action is less pronounced. The reverse is equally possible. Furthermore, an outgrowth of new myelin tubes can be obtained when the immobilizing calcium solution is rendered more alkaline, whereas acidulation causes a further shrinkage of the bodies. These reactions may be reversed several times if the salt concentration is kept within a range corresponding to the one tolerated by living cells.

The apparent antagonistic effect of bivalent and monovalent cations upon the swelling of lecithin bodies is actually of a merely quantitative nature since, at sufficiently high concentrations, both achieve complete desolvation. In consequence of the competition of the different cations for a combination with the phosphoric acid groups, lecithin is capable of exhibiting a moderate rate of myelinitic growth in vertebrate Ringer, although the outgrowth is much reduced if the proportion of calcium ions contained in this solution is applied in the absence of the growth-promoting monovalent cations. The repressive effect of high salt concentrations upon the swelling of phosphatide vesicles may account for the absence of a contractile vacuole in many marine Protozoa, and for its disappearance when a fresh-water protozoan is transferred to sea water. As to the readiness with which cations are exchanged between living cells and their environment, reference may be made to the reviews which appeared in the *Symposium on Quantitative Biology*, volume 8 (1940).

Permeability of Myelin Bodies. Swelling and shrinkage of a lecithin vesicle would be impossible if its external wall were not readily permeable to water. Apart from water, a great variety of solutes have been observed to enter the vesicles. The criteria used were internal precipitations occurring under the influence of the permeated substance, or changes of color of the inner fluid.

It is clear from the preceding paragraphs that the different metal ions mentioned above have no difficulty in passing through the interfacial lecithin membrane, their rate of penetration being measured in seconds rather than minutes. Shrinkage and internal precipitations occur in solutions of histone, egg albumen, and various amino acids (Holtfreter, 1946c). Hydrogen and hydroxyl ions may equally well enter into a myelin vesicle. This can be demonstrated by staining the vesicle in a solution containing an adequate pH indicator dye and observing the change of color occurring inside the vesicle after an acid or a base has been added to the solution. Like the histological "acid dyes" (fast green, eosin, orange G, acid fuchsin), the customary indicator dyes were found to enter the vesicles without producing noticeable morphological changes. These dyes do not precipitate with the lipids, hence are not "stored," and diffuse out again when the vesicles are returned to pure water. They act, therefore, quite differently from "basic dyes" (neutral red, Nile blue, toluidin blue, basic fuchsin, methylene blue), which are taken up from very dilute solutions, become highly concentrated both in the wall and in the internal fluid of the shrinking myelin bodies, and cannot be washed out again because they combine with the phosphatide molecules which become dehydrated. Comparable differences of staining affinity for the two groups of dyes are observed in living cells and their vacuoles and other lipid structures. It must be added that the acidity of rancid lecithin changes the color of the permeated indicator dyes according to their pH sensitivity. The pH of the fluid within the myelin bodies used was found to be about 4.8.

The volume changes of the vesicles occurring in the different solutions discussed above cannot be satisfactorily explained by differences of osmotic pressure between the internal fluid and the external suspension medium, but are primarily caused by varying states of condensation of the limiting membrane. Strong concentrations of desolvating chemicals can be observed to cause a shrinkage of successive surface layers which slide over and peel off from the subjacent, less desolvated layers of the multimolecular wall of myelin bodies.* While contracting, the layers exert a marked pressure upon the enclosed liquid, which may escape explosively through a thinned-out and bursting portion of the wall. These observations may account for the fact that the cell membrane of living cells may burst not only in alkaline or hypotonic media, but also under the influence of histone, or basic dyes, which actually desolvate the membrane. It would appear that, analogous to the myelin vesicles, experimentally induced volume changes of embryonic cells are initiated by physico-chemical changes of the cell membrane rather than by osmotically controlled variations of turgor of the inner cytoplasm.

Models of Pseudopods and of Amoeboid Movements. When a fully extended myelin tube is dried on glass and subsequently covered with water, it sprouts out into numerous thin-walled tubules and vesicles (FIGURE 22). This indicates that through desiccation the homogeneous molecular layers of the tubular wall are disarranged and become oriented in new growth patterns when refurnished with water. Corresponding configurations are obtained when a swollen myelin vesicle is at first partly dehydrated by acid and then exposed to alkali (FIGURE 21). Again, the previously homogeneous surface membrane of the vesicle becomes differentiated into regions which vary in their readiness to expand when the conditions of swelling are restored. This results in the outgrowth of

---

* The occurrence of sliding and churning movements within the wall of a myelin body, or of translocations of particles attached to its surface, can be explained by the common assumption that the superimposed bimolecular lipid lamellae are separated by "lubricating" layers of water and may expand and contract independently. Such movements may be wrongly interpreted as indicating a liquid state of the entire wall. Similar translocations have been observed in particles attached to the cell surface of migrating amoebae (Schaeffer, 1920) and of dividing sea urchin eggs (Dan, Yamagita, and Sugiyama, 1937), while churning movements occurred in the plasma membrane of denuded sea urchin eggs which had been touched by a glass needle (Chambers, 1938). It is still a matter of argument how many molecular lipid layers may be present in living cell membranes (Schmitt, Bear, and Ponder, 1936, 1938; Danielli, 1942). If this structure should consist of several molecular layers, then the flowing movements observed might involve a superficial layer having only paracrystalline structure and being, therefore, not in a "liquid" state.

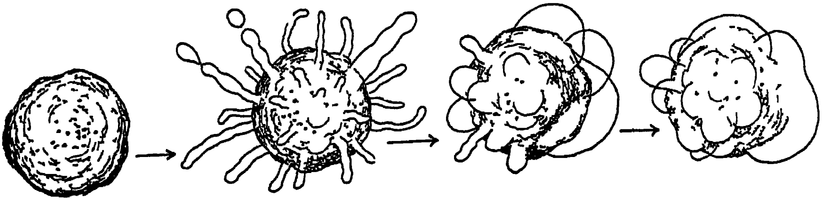
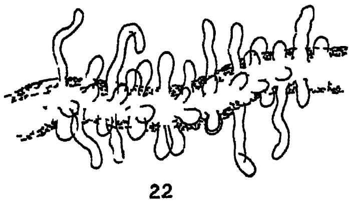


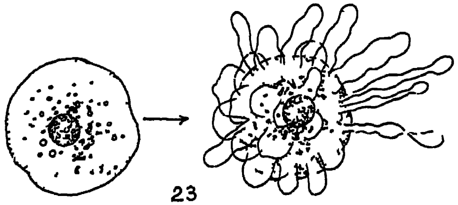
FIGURE 21 Myelin vesicle already shrunk in acid, forms at first tubular, then lobose processes when subsequently exposed to alkali

numerous processes, the size and shape of which resemble the hyaline pseudopods which are produced in living cells by the application of alkali (compare FIGURES 21 and 9). With a further raising of the external pH, larger areas of the surface film become expanded and the tubular processes change into voluminous bulges which, by way of local and temporal variations in the rate of solvation, constantly change their outlines. Finally, some portions of the swollen periphery may become pinched off in the form of vesicles (FIGURE 21).

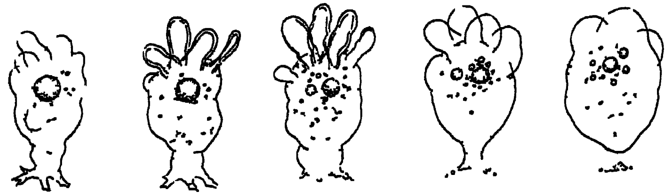
Such form changes can be induced by the successive application of many other antagonistic agents which need not have more in common than that the first one has a comparatively stronger desolvating effect than the second one. Instead of using HCl or other acids, one may partly desolvate the myelin bodies by the application of bi- or multivalent metal ions, of basic dyes, or of certain proteins (histone, egg albumen). The subsequent local swellings can be induced by the arbitrary application of alkali, of monovalent cations, or simply by diluting the solution. The process is illustrated in FIGURE 23, which shows a myelin body of the size of a gastrula cell which has first undergone shrinkage



22



23



24

FIGURE 22 Numerous filial buds growing out from a myelin tube that has been dried and subsequently wetted

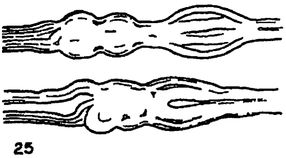
FIGURE 23 Myelin body, partly dehydrated by CaCl$_2$, exhibits tubular outgrowth after the addition of NaCl to the solution

FIGURE 24 Polar tubular outgrowth in a myelin body containing fatty acid droplets

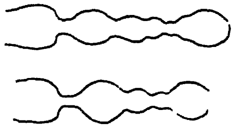
through the action of $CaCl_2$, and subsequently expanded into tubular and lobose processes when NaCl was added to the solution.

In addition to the external variables, local differences in the composition of the lipid body may cause the swelling process to become differentiated into a polar growth pattern. This is shown in the example of FIGURE 24, where regional differences of myelinitic activity are evidently correlated to the distribution of droplets of fatty acid which have accumulated in the upper region of the pear-shaped body. The presence of free fatty acids results from the advanced state of oxidation of the lecithin employed. While in fully expanded myelin vesicles the fatty acid particles are, in general, invisibly dispersed, any dehydrating chemical causes them to coalesce into granules and droplets of increasing diameter, which may be finally expelled into the external medium. In the present case, the myelin vesicle had been dehydrated by a strong solution of chrysoidin. The dye rendered the body adhesive to glass and stained it yellow, whereas the coalescing fat droplets stained red. Following the addition of water to the dye solution, the body performed a succession of irregular expansions and contractions, while the surface membrane in the region of the large fat droplet projected a series of pseudopodia-like processes which constantly changed their shape and eventually fused into larger bulges (FIGURE 24). With a progressive swelling of the whole body, the "foot" detached itself from the glass and the fat droplet broke up again into smaller particles which became dispersed in the internal fluid.

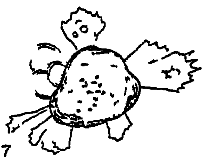
Differential swelling of an individual lecithin body can be obtained, furthermore, by introducing into it a particle of a dehydrating substance, such as a basic dye When the body is immersed in water, the extent of swelling of the different regions increases with their distance from the incorporated particle. However, even in the absence of formed in-



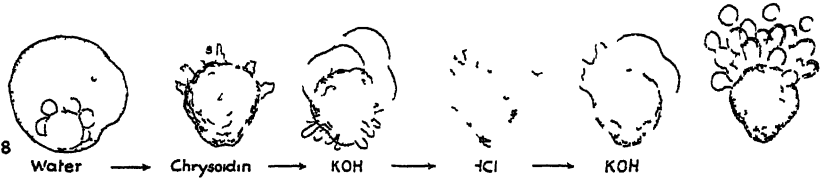
25



26



27



28

FIGURE 25 Changing pattern of varicosities in a water-immersed thick-walled myelin tube

FIGURE 26 Two phases of a constriction pattern in a myelin tube

FIGURE 27 Membranous myelin processes spreading on glass, following exposure to a strong Nile blue solution

FIGURE 28 Alternate shrinkage and swelling of circumscribed regions of a myelin body

clusions, myelin formations may exhibit regional differences of swelling, indicating that the walls themselves may be of an inhomogeneous composition. Frequently, a myelin tube consists of abruptly set-off segments, each having its own caliber and swelling tendency (FIGURE 25). In other instances, the wall appears to be continuous and of equal diameter, yet the body may grow out into side branches, or show along its length an alternation of bulges and constrictions resembling somewhat those observed in pseudopods (FIGURE 26). Although a regular peristaltic progression of the constrictions did not occur, it was frequently observed that terminal portions became pinched off from a myelin tube. As in living cells, fragmentation could be induced by both hydrating and dehydrating agents.

An imitation of "spontaneous" amoeboid movements was observed in the example of FIGURE 29. This myelin formation, while suspended in a dilute solution of alloxan, shrank somewhat and its surface exhibited for a longer period local bulges which were leveled out again and could reappear in the same or in another region of the body. Similar movements of a slightly periodic nature were observed in myelin formations which had been mildly dehydrated by a solution of Nile blue (Holtfreter, 1946c, Figure 16). It may be assumed that the movements were brought about by the antagonism between water and the desolvating agent, the two competing for a combination with the phosphatide molecules and replacing each other alternatingly.

Laminar Processes in Myelin Bodies. Fully hydrated myelin bodies which are non-adhesive to glass, to each other, and to the cell surface, will stick to and eventually spread over these surfaces when sufficiently dehydrated by any of the following agents: alcohol, acids, metal ions, basic dyes, or basic proteins. FIGURE 28 shows the successive stages of transformation produced in a myelin vesicle by the alternate application of hydrating and dehydrating agents. It will be noticed that, when the previously non-attached body shrinks and flattens over the glass surface, its periphery forms a serrated margin. In consequence of dehydration, the lipid components become arranged in new patterns, forming vacuoles, rings, and reticular ridges which can be interpreted in terms of coacervation (Bungenberg de Jong, 1932, 1935). These flowing surface patterns resemble the ruffles and networks which occur in the margin of migrating cells and, in a coarser form, in hyaline fragments that have been transformed into an adhesive film by way of dehydrating fixatives (compare FIGURES 20a, 27, 30c, and 10).

Locomotion of Lipid Bodies. It is well known that, when an oil droplet floating at the water surface is combined with a surface tension lowering substance, such as alkali or lecithin, its spreading is locally increased, resulting in the formation of mobile lobes which may become detached from the periphery. The following experiment illustrates how the local

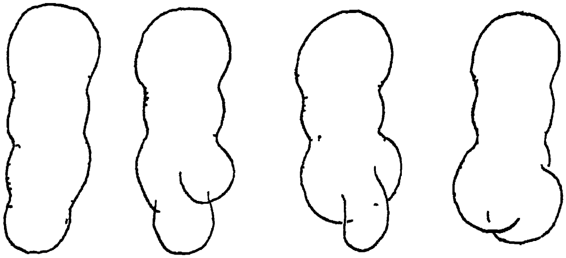


FIGURE 29. Form changes of a myelin vesicle in a solution of alloxan.

FIGURE 30. Myelin bodies showing hyaline bulges produced by alkali (a, b), or a flattened and adhesive margin with surface wrinkles, caused by acid treatment (c).

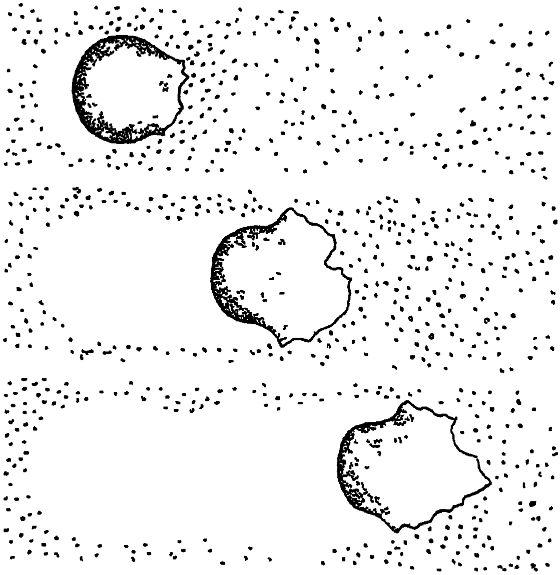


FIGURE 31. A droplet of oleic acid moving through a film of cell debris

incorporation of protoplasmic constituents may induce a unidirectional locomotion in floating oil droplets.

When an embryonic amphibian cell is brought to the air-water interface, it bursts and releases its content, which spreads as a thin film dotted with fat droplets. The latter originate from the lipochondria, which consist predominantly of neutral fat, phosphatides, and cholesterol (Fauré-Fremiet and de Streel, 1921; Holtfreter, 1946c). When subsequently a small drop of oleic acid is injected beneath this film, it attracts the adjacent lipid components of the film, which are then vigorously absorbed. That side of the drop which is incorporating the surface tension-lowering substances flattens and advances in irregular to-and-fro movements, while the opposite side remains spherical and immobile (FIGURE 31). Thus, the drop assumes the polar shape and kinetic activity of a migrating cell. It moves forward by "eating" its way through the film, leaving behind a trail free of visible particles. Eventually, the whole drop flattens out and becomes quiescent.

Membrane Formation in Relation to Gibbs-Thompson's Law. It is not claimed that the preceding experiment is an ideal illustration of the mechanism of amoeboid movements. However, it shows that the protoplasm of the amphibian cell contains substances capable of inducing surface movements when incorporated into an interfacial layer of another lipid. Previous experiments have shown that the lipid constituents of the lipochondria may combine with the dissolved protein of the yolk to form vesicles bounded by a film which has properties similar to those of a cell membrane (Holtfreter, 1947a). Apart from the granular cell inclusions, the naked cytoplasm of the amphibian egg, when exposed to any kind of hypotonic salt solutions, is capable of elaborating membranous structures in the form of vacuoles or vesicles. It may be assumed that, in normal embryos, the constituents of the cell membrane are recruited from the lipoproteins of the cytoplasm and that the increase of cell surfaces occurring during development is associated with an incorporation of more of these substances into the outer membrane.

Like other surface active substances, phosphatides, obeying Gibbs-Thompson's law, tend to accumulate at interfaces. This can be demonstrated by stirring phosphatide bodies into a water-immersed drop of triolein (Holtfreter, 1946c, Figure 6). The bodies move centrifugally toward the surface of the drop, where they spread as an interfacial film showing the birefringence characteristic of radially oriented lipid molecules. It is perhaps due to the same principle that the pigment granules of the egg tend to accumulate near the egg periphery and along other interfaces, irrespective of their own specific gravity. Other possible examples of this principle are the centrifugal movements of the nuclei in centrolecithal eggs, and the migration of ectoderm cells from deeper layers into the surface of the amphibian embryo. Once the substances of various size and composition have attained a cortical position, they will react with each other and with the constituents of the external phase to form compounds and structures which are absent in deeper layers of the system. The formation of the vitelline membrane and the external coat may be manifestations of such specific interfacial conditions. Furthermore, the pronounced spreading tendency of the ectoderm as well as the proximo-distal polarity of all cells developing from the egg surface are possibly expressions of inside-outside differentials of the egg correlated with the Gibbs-Thompson phenomenon.

In colloidal mixtures, phosphatides not only tend to accumulate at the outer interface, but they may create new interfaces in the form of internal vacuoles. The physico-chemical mechanism and the cytological implications of this phenomenon have been discussed by Bungenberg de Jong (1932, 1935). In connection with this type of coacervation, another phenomenon may be recorded. It was observed that, when the composite lipid extract from lipochondria was spread on glass as a thin layer and covered by a Nile blue solution, it became parceled up into cell-like compartments, differing one from another in their lipid composition, each unit being surrounded by a film of phosphatides (Holtfreter, 1946c). This process recalls the neoformation of cell boundaries in polynuclear syncytia, such as occurs in slime molds, or in the blastoderm of various egg types.

General Conclusions

The wide range of similarities existing between the living cell membrane and artificial phosphatide structures is a challenge to base an interpretation of the functions of the cell membrane upon some of the concepts gained from lipid models. In both the cell and myelin bodies, there was an interrelationship between the phenomena of adhesion, swelling, and motility. When embryonic cells were exposed to conditions which favor the swelling of myelin bodies (high pH, lack of calcium, low salt concentration), they underwent the following changes: loss of adhesiveness, increase of the amount of ectoplasmic fluid, rounding-up of filiform into lobose pseudopods, spreading of amoeboid motility over previously quiescent areas of the cell surface, acceleration of the movements, increase in the rate of endoplasmic eruptions. These phenomena could be reversed by the application of agents such as low pH, hypertonicity, or calcium, which produced shrinkage and adhesiveness in myelin bodies. However, prolonged exposure of the cell to the liquefying agents caused irreversible changes. Amoeboid motility faded out, the endoplasm became dispersed throughout the spherically swollen cell, the cell membrane became flabby and permeable to cell inclusions of increasing size, and finally either coagulated or disintegrated entirely.

Evidently, these phenomena of cytolysis are primarily caused by irreversible changes in the non-lipid constituents of the cell, particularly by a breakdown of protein compounds. An analogy with myelin vesicles exists only in so far as their walls likewise become thinned out and may burst when exposed to alkali. It is probably because of the comparatively greater complexity of the cell structures that their irreversible changes occurring under extreme culture conditions find no true analogy in the behavior of lecithin vesicles. Therefore, a comparison of the two systems should be confined to the conditions which maintain cellular viability.

Polar Structure of the Cell Membrane. The kinetic behavior of isolated amphibian cells showed that their capability of assuming specific shapes develops gradually, but that, independent of their prospective significance, the neurula cells already possess an antero-posterior polarity. This primordial organization was manifested in polar differences of adhesiveness and amoeboid activity, as well as in the tendency of the cells to elongate reversibly into cylindrical bodies. While undergoing further differentiation, the individual cells acquired new and more elaborate patterns of adhesiveness and motility, which determined their direction of locomotion and their tissue-specific shapes. This process was associated with an intracellular digestion of the embryonic lipoprotein inclusions and, apparently, with a certain amount of dehydration.

Observational evidence pointed to the conclusion that the polar organization of the embryonic cell is primarily controlled by the cell membrane. The facts that the posterior pole always represents the originally distal side of the cell, and that the uncoated endoderm cells from the interior of the gastrula do not clearly reveal a polarity, seem to indicate that the polarity originates from the inside-outside gradient of the egg mentioned above. The polar characteristics are partly due to the fact that the posterior cell pole retains some of the properties of a coated cell surface, such as reduced permeability and adhesiveness, pronounced contractility, and the absence of a fluid layer between plasmagel and cell membrane.

However, these considerations do not yet explain the capability of the cell to form cylindrical or actinian-shaped bodies. This faculty must arise from axial differences of structure encompassing the whole cell surface. It has been shown above that, as a consequence of induced local differences of swelling power, the originally homogeneous surface layer of a myelin body may become heterogeneous and exhibit localized tubular growth. Accordingly, it may be assumed that, because of regional differences in the composition of the cell membrane, its response to solvating reagents increases gradually from the posterior to the anterior cell pole, and that in later stages further local differences of condensation of the membrane are developed.

In ectoderm cells, the faculty of adopting a specific shape arises at the time of their cytological determination, suggesting that the process of induction liberates certain compounds, perhaps proteins, which become integrated formative elements of the outer membrane. There is no fundamental difficulty in explaining the various forms of amoeboid movement in terms of solvation of oriented lipid layers of the cell surface. However, it is conceivable that tangentially arranged fibrillar protein constituents of the cell membrane not only determine the elasticity of the cell surface (Harvey and Danielli, 1936) and contribute in differentiating the shape of the cell, but that in cooperation with the lipids they are directly responsible for amoeboid movements. Analogous to myosin micelles, proteins contained in the cell membrane may undergo reversible molecular contractions and extensions. The fact, however, that, in contrast to the lipid models, no protein models have as yet been devised which imitate in any way the various functions of the cell membrane, speaks in favor of the view that the essential kinetic elements of the membrane are the lipids rather than proteins.

Cellular Adhesiveness. It appears that the most important general factor determining, simultaneously, cellular adhesion, permeability, and motility is the state of solvation of the cell membrane. The physiological importance of the mono- and bivalent cations in controlling these phenomena is possibly derived from their readiness of substituting each other in protoplasmic compounds. Adhesion and aggregation of embryonic amphibian cells can be brought about by calcium ions in the ambient salt solution, provided the latter is isotonic and not too alkaline. The bivalent cations seem to operate according to the "zipper mechanism" of Schmitt (1941), by desolvating and pulling together the contacting cell surfaces. This concept presupposes a preponderance of acidic groups, possibly phosphoric acids, at the surface of the cell membrane, an assumption which finds support in the cataphoretic behavior of some cells, particularly of erythrocytes.

While in early embryonic stages only calcium and hydrogen ions seem to be required in order to hold the cells together, an intercellular adhesive matrix, presumably of a protein nature, is reported to exist in between the cells of differentiated epithelia (Gray, 1926; Chambers, 1940). Further insight into the nature of such cementing substances may be gained from investigations of the kind which Chargaff and co-workers (1944) made on the thromboplastic lipoproteins and their function in the process of blood clotting.

The cells in the interior of early amphibian embryos are not firmly aggregated. This is, perhaps, a prerequisite for the execution of the morphogenetic movements which occur during this period. The lack of cellular adhesion apparently arises from the absence, or inactivation, of agglutinating substances in the body fluid, since the cells of the different germ layers will indiscriminately unite with each other when exposed to a balanced salt solution.

With progressive differentiation, there arise cell-specific differences of adhesiveness which are reflected in the display of histotypical patterns of aggregation, disaggregation, migration, and recombination of the various cell strains (Holtfreter, 1939, 1944). It is not known whether these manifestations of a selective adhesiveness result from a molecular lock-and-key mechanism of the naked cell surfaces (Weiss, 1941) or from the interference of specific cementing substances, which may either exudate from the contacting cells themselves or be furnished by the external fluid. Weiss (1947), in a recent discussion of the significance of cellular "affinities" and "disaffinities" for the elaboration of growth patterns, has drawn attention to the similarities which seem to exist between these phenomena and serological reactions. However, aggregations are possible between cells or myelin bodies, the membranes of which may be assumed to have an identical molecular configuration. Nor is it likely that the adhesions occurring between cells and other substances, such as glass, oil droplets, fibrin, etc., are caused by the presence of complementary sets of polar groups at the contact surfaces, comparable to those which are thought to determine antigen-antibody reactions. The ease with which cellular adhesions can be broken up and restored, and the fact that cells migrate in spite of adhesion, indicate that the bonds of attachment are very labile. When observing the gliding movements performed by the surfaces of contacting cells, one gains the impression that the attraction forces operate across an intercalated cushion of water (Holtfreter, 1946b), suggesting that the merely temporary adhesion of migrating cells may be controlled by long-distance forces of the kind which produce the tactoid formation in certain colloidal sols.

Amoeboid Movements. The general rule that cellular adhesiveness decreases with the swelling of the cell likewise applies to the different regions of the individual cell, for the surface of an expanded hyaline bulge is non-adhesive in contrast to the more contracted region of the cell. It seems, therefore, that both non-adhesiveness and expansion of the cell membrane involve a state of increased solvation of this structure, whereas a more contracted and adhesive condition of the membrane would indicate the loss of bound water. The amoeboid movements described above consist essentially of two phenomena: (*1*) periodic expansions and contractions of certain areas of the cell membrane; and (*2*) a more or less regular propagation of these alternate states of activity along the antero-posterior axis of the cell. The first phenomenon appears to be more readily accessible to an interpretation than the second one. With Fauré-Fremiet (1929), it may be assumed that, analogous to the phosphatide models, the movements reflect periodic variations in the spacing of the radially arranged lipid molecules contained in the cell membrane. X-ray analysis and polarization-optical studies of phosphatides indicate that variations in the water content of these structures not only alter the distance between the bimolecular lipid leaflets but also affect the interchain packing of the molecules (Palmer and Schmitt, 1941). Both in cells and phosphatide models reversible expansions and contractions of the external membrane can be brought about by the alternate application of a great variety of respectively solvating and desolvating agents. However, it is a matter for conjecture what are the intracellular processes that produce the movements in the absence of external stimuli. Probably an indispensable regulative role in the kinetic function of the cell membrane can be ascribed to the antagonistically active hydrogen ions and bivalent metal ions on the one hand, and the hydroxyl ions and monovalent cations on the other. In addition, a reversible interaction between the phosphatide molecules and the cholesterol and proteins present in the cell membrane may determine the extent of expansion of this layer. Bear, Palmer, and Schmitt (1941) have shown that, in the presence of water, the previously incompatible molecules of various lipids (phosphatides, cerebrosides, and cholesterol) will form a homogeneous mixed phase where the different constituents become lined up side by side, showing a single identity period. Monolayers of lecithin undergo a closer molecular packing and, hence, a contraction, when cholesterol is intercalated between the lecithin molecules (Leathes, 1925). In their attempt at interpreting the structure and function of the cell membrane on the basis of coacervate models, Bungenberg de Jong and co-workers (1932, 1935) have pointed to the desolvating and tightening effect of cholesterol, triolein, and oleic acid upon phosphatide structures, this process being strongly influenced by the presence of electrolytes.

As to the effect of proteins upon the packing of oriented phosphatide layers, reference has already been made to the pronounced desolvating action of basic proteins (histone, albumen). However, since the resulting lipoprotein compounds are practically water-insoluble, it appears improbable that strongly basic proteins are engaged in amoeboid surface contraction, which is a reversible phenomenon. More adequate models of the molecular interactions in lipoprotein membranes seem to be represented by the complex monolayers studied by Schulman and Rideal (1937). These authors showed that certain mixtures of cholesterol and wheat gliadin form a composite liquid film on the air-water interface, which gelates on compression. When the pressure is further increased, the film liquefies suddenly, because the gliadin is driven from the surface film into the underlying water. On decompression, the protein molecules reenter the cholesterol film, which expands. This gelation-liquefaction process is reversible several times. It is very sensitive to pH and to the ratio of the lipid-protein concentration. The applicability of these studies to cytological problems is underlined by the observation that cytolytic agents, such as fatty acids or soaps, rapidly penetrate and disperse the artificial lipoprotein films, while such agglutinating substances as tannic acid and gallic acid link the protein molecules together into a hydrophobic "skin" which resists the penetration of soaps.

On the basis of the above considerations, amoeboid movements may be conceived as resulting from localized and alternate states of solvation of the lipoprotein envelope of the cell, these changes being caused by reversible interactions between the radially arranged lipid molecules and other compounds, possibly proteins. The efficiency of this mechanism depends upon a balanced ionic atmosphere. The process may be somewhat analogous to the reversible combinations occurring between numerous enzymes, or pigments, and their carriers, where again lipids and proteins are the most universally present constituents of the systems (Needham, 1942, p. 206; Chargaff, 1944). As to the periodicity and propagation of the surface movements, electric phenomena, comparable to those which accompany the nerve impulse, appear to be involved. Hubbard and Rothschild (1939) made the interesting observation that both the unfertilized and segmented eggs of the trout exhibit rhythmical changes of impedance, of the frequency of about 1.5 per minute. The changes are interpreted as possibly caused by a thickening or thinning of the protoplasmic membrane, involving "a change in the ability of polar or oriented molecules to rotate according to the sense of the applied current."

Summary

Based mainly upon observations on amphibian material, the attempt has been made to show that many embryological phenomena may be better understood if we take into consideration the properties and functions of the interfacial membranes which separate the cells from each other and from the external medium. While all cells are furnished with a living plasma membrane, the periphery of the amphibian egg, and the epithelia deriving from it, possess an additional covering in the form of a coat, which resembles in many ways the hyaline layer in echinoderm eggs. The coat, though not considered to be a living and indispensable part of the egg, plays an important role in determining the viability of the embryo under various environmental conditions, and in controlling, but not causing, the morphogenetic movements of gastrulation and neurulation. This structure seems to consist predominantly of protein compounds, containing calcium, which with progressive differentiation become more densified and less soluble.

The cells derived from the egg periphery possess a proximo-distal polarity which is expressed in polar differences of adhesiveness and amoeboid activity, and in the tendency of these cells to stretch themselves reversibly along an antero-posterior axis. These phenomena are thought to result from regional differences in the composition of the cell membrane, which in turn reflect an inside-outside gradient of the egg connected with the Gibbs-Thompson effect. Subsequent form changes and directed locomotions of the cells are decisively influenced by their inherent axial polarity.

The kinetic behavior of isolated embryonic cells under various environmental conditions indicates that cellular form changes and locomotion result primarily from rhythmic expansions and contractions of the cell membrane. The endoplasmic core, which is more or less separated from the outer membrane by a layer of ectoplasmic fluid, may undergo cyclic sol-gel formations which are, however, not the cause of amoeboid

movements. A cell becomes fragmented if a constriction wave passing over the cell surface becomes stationary and cuts progressively deeper into the cell body. This phenomenon is brought into relation to normal cytoplasmic division.

Local differences of cellular adhesiveness and amoeboid activity, arising in the course of development, are regarded as the main factors which transform the primitive cell into the specific shapes characteristic of the different cell strains. General features of this differentiation process are the replacement of lobose pseudopods by filiform and lamellar processes, attachment of the cell membrane to the plasmagel over the major surface area of the cell, and progressive confinement of amoeboid motility to the tips of the extended pseudopods, which remain free of endoplasm. This process seems to involve a certain amount of dehydration since, on exposure to liquefying solutions, half-way differentiated cells may readopt the appearances and kinetics of earlier developmental stages. After the cell-specific shape has been established, it may become consolidated by the elaboration of an inner cytoskeleton.

The direction of cellular migration, and the histotypical groupings and regroupings exhibited by the various types of cells in a developing organism, appear to be controlled by a selective adhesiveness of the cell membrane, which varies with the developmental stage and with the kind of cells involved. Cellular adhesiveness depends both on the chemical constitution of the contacting cell surfaces and on the composition of the immersion fluid. From the observed antagonistic effects of hydrating and dehydrating agents upon cellular adhesion, it may be concluded that the most universal and essential factor determining adhesion is the degree of solvation of the cell membrane. A prerequisite for cellular aggregation appears to be a certain concentration of hydrogen and calcium ions in the ambient solution.

The attempt is made to interpret the kinetic functions of the cell membrane on the basis of data on the chemical and physical properties of this structure. According to the evidence available, the limiting plasma film of eggs and somatic cells consists predominantly of alternating lamellae of oriented protein and lipid molecules, the latter being chiefly represented by phospholipids. In support of this concept, it is shown that many of the features pertaining to cellular permeability, adhesion, and surface movements can be imitated in models of hydrated phosphatide bodies which are subjected to various environmental conditions. The surprising similarities in the physico-chemical behavior of cells and myelin formations suggest that amoeboid movements result from alternate states of packing of the oriented lipid molecules of the cell membrane. These changes seem to involve reversible variations in the extent of solvation of the lipid leaflets, brought about by the action of electrolytes and, perhaps, proteins and other compounds. The wave-like propagation of the alternate states of film condensation along the cell surface recalls the impulses traveling along a nerve fiber. The above concept would facilitate an understanding of the fact that cellular motility, adhesion, and permeability are usually correlated phenomena and that all three are affected when the cell is acted upon by agents which interfere with the physico-chemical conditions of the cell membrane.

Bibliography

Angerer, C. A. 1936. The effects of mechanical agitation on the relative viscosity of *Amoeba proteus.* J. Cell. &. Comp. Physiol. 8: 329.

Battle, H. K. 1944. Effects of dropping on the subsequent hatching of teleostean ova. J. Fish. Res. Board Canada 6: 252.

Bear, R. S., R. J. Palmer, & F. O. Schmitt. 1941. X-ray diffraction studies of nerve lipides. J. Cell. & Comp. Physiol. 17: 355.

Briggs, R. W. 1941. The development of abnormal growths in *Rana pipiens* embryos following delayed fertilization. Anat. Rec. 81: 121.

Bungenberg de Jong, H. G. 1932. Die Koacervation und ihre Bedeutung fur die Biologie. Protopl. 15: 150.

Bungenberg de Jong, H. G., & J. Bonner. 1935. Phosphatide auto-complex coacervates as ionic systems and their relation to the protoplasmic membrane. Protopl. 24: 198.

Bungenberg de Jong, H. G., & G. G. P. Saubert. 1937. Fortschritte zum Thema der Modelle der Protoplasmamembran. Protopl. 28: 352.

Chalkley, H. W. 1935. The mechanism of cytoplasmic fission in *Amoeba proteus.* Protopl. 24: 607.

Chambers, R. 1924. The Physical Structure of Protoplasm as Determined by Microdissection and Injection. In: Cowdry, E. V. *General Cytology.* Univ. of Chicago Press.

1938. Structural and kinetic aspects of cell division. J. Cell. & Comp. Physiol. 12: 149.

1940. The relation of extraneous coats to the organization and permeability of cellular membranes. Symp. Quant. Biol. 8: 144.

Chargaff, E. 1944. Lipoproteins. Adv. Protein Chemistry 1.

Chinn, P. 1938. Polarization optical studies of the structure of nerve cells. J. Cell. & Comp. Physiol. 12: 505.

Crile, G., M. Telkes, & A. F. Rowland. 1932. Autosynthetic cells. Protopl. 15: 337.

Dan, K., T. Yamagita, & M. Sugiyama. 1937. Behavior of the cell surface during cleavage. Protopl. 38: 66.

Danielli, J. F. 1936. Some properties of lipoid films in relation to the structure of the plasma membrane. J. Cell. & Comp. Physiol. 7: 393.

1942. The Cell Surface and Cell Physiology. In: G. Bourne. Cytology and Cell Physiology: 68–98. Clarendon Press. Oxford.

1945. Reactions at interfaces and their significance in biology. Nature 156: 168.

Dareste, C. 1891. Recherches sur la production artificielle des monstrosités ou essai de tératogénie expérimentale. C. Reinwald & Cie. Paris.

Dawson, H., & J. F. Danielli. 1938. Studies on the permeability of erythrocytes. Biochem. J. 32: 991.

Erlanger, R. 1897. Über die Morphologie der Zelle und den Mechanismus der Zellteilung. Zool. Centrbl. 4.

Erickson, B. N., H. H. Williams, S. S. Bernstein, J. Arvin, R. J. Jones, & I. G. Macy. 1938. The lipid distribution of posthemolytic residue or stroma of erythrocytes. J. Biochem. 122: 515.

Fankhauser, G. 1934. Cytological studies on egg fragments of the salamander *Triton.* IV. J. Exp. Zool. 67: 349.

Fauré-Fremiet, E. 1925. *La Cinétique du Développement.* Presse Universitaire. Paris.

1929. Caractères physico-chimiques des choanoleucocytes de quelques invertébrés. Protopl. 6: 521.

1930. The kinetics of living matter. Trans. Faraday Soc. 26: 779.

FAURÉ-FREMIET, E., & MLLE. DUVIVIER DE STREEL. 1921. Les constituents chimiques de l'oeuf et leur rôle dans le développement embryonnaire chez la grenouille rousse (*Rana temporaria*). Bull. Soc. Chim. Biol. 3: 476.

FREY-WYSSLING, A. 1938. Submikroskopische Morphologie des Protoplasmas und seiner Derivate. Protopl. Monogr. Borntraeger, Berlin.

GRAY, J. 1926. The properties of an intercellular matrix and its relation to electrolytes. Brit. J. Exp. Biol. 3: 167.

HARVEY, E. B. 1936. Parthenogenetic merogony or cleavage without nuclei in *Arbacia punctulata*. Biol. Bull. 71: 101.

1938. Parthenogenetic merogony or development without nuclei of the eggs of sea urchins from Naples. Biol. Bull. 75: 170.

1939. Development of half-eggs of *Chaetopterus pergamentaceus* with special reference to parthenogenetic merogony. Biol. Bull. 76: 384.

HARVEY, E. N. 1912. A new type of artificial cell suitable for permeability and other biochemical studies. Biochem. Bull. 2: 50.

1936. The properties of elastic membranes with special reference to the cell surface. J. Cell. & Comp. Physiol. 8: 251.

HARVEY, E. N., & J. F. DANIELLI. 1936. The elasticity of thin films in relation to the cell surface. J. Cell. & Comp. Physiol. 8: 31.

1938. Properties of the cell surface. Biol. Rev. 13: 319.

HERRERA, A. 1932. Rectifications historiques à propos des cellules autosynthétiques de M. le Dr. G. W. Crile. Protopl. 15: 361.

HERWERDEN, M. A. VAN. 1933. Erscheinungen an elementaren fibrillären Membranen von mesomorphem Charakter. Protopl. 19: 312.

HOBSON, L. B. 1941. On the ultrastructure of the neural plate and tube of the early chick embryo. J. Exp. Zool. 88: 107.

HOLTFRETER, J. 1933. Die totale Exogastrulation, eine Selbstablösung des Ektoderms vom Entomesoderm. Roux' Arch. 129: 669.

1939. Gewebeaffinität, ein Mittel der embryonalen Formbildung. Arch. exp. Zellf. 23: 169.

1943a. Properties and functions of the surface coat in amphibian embryos. J. Exp. Zool. 93: 251.

1943b. A study of the mechanics of gastrulation. I. J. Exp. Zool. 94: 261.

1944a. A study of the mechanics of gastrulation. II. J. Exp. Zool. 95: 171.

1944b. Neural differentiation of ectoderm through exposure to saline solution. J. Exp. Zool. 95: 307.

1946a. Structure, motility, and locomotion in isolated embryonic amphibian cells. J. Morphol. 79: 27.

1946b. Observations on the migration, aggregation, and phagocytosis of embryonic cells. J. Morphol. 80: 25.

1946c. Experiments on the formed inclusions of the amphibian egg. II. J. Exp. Zool. 102: 51.

1947a. Experiments on the formed inclusions of the amphibian egg. III. J. Exp. Zool. 103: 81.

1947b. Changes of structure and the kinetics of differentiating embryonic cells. J. Morphol. 90: 57.

1947c.. Morphogenesis, crenation and cytolytic reactions of the erythrocytes of amphibians. J. Morphol. 80: 345.

HUBBARD, M. I., & LORD ROTHSCHILD. 1939. Spontaneous rhythmical impedance changes in the trout's egg. Proc. Roy. Soc. London B 127: 510.

HYMAN, L. H. 1917. Metabolic gradients in *Amoeba* and their relation to the mechanics of amoeboid movement. J. Exp. Zool. 24: 55.

JENKINSON, J. W. 1906. On the effect of certain solutions upon the development of the frog's egg. Arch. Entw.-Mech. 21: 367.

JENNINGS, H. S. 1904. Contributions to the study of the behavior of lower organisms. Publ. Carnegie Inst. Wash.

JOLLOS, V., & T. PETERFI. 1923. Furchung von Axolotleiern ohne Beteiligung des Kernes. Biol. Zentralbl. 43: 286.

JUST, E. 1922. Studies on cell division. Am. J. Physiol. 61: 505.

1938. *The Biology of the Cell Surface.* Blakiston. Philadelphia.

KÖLSCH, K. 1902. Untersuchungen über die Zerfliessungserscheinungen der Ciliaten Infusorien. Zool. Jahrb. (Anat.) 16: 173.

LANDAUER, W., & L. BAUMANN. 1943. Rumplessness of chicken embryos produced by mechanical shaking of eggs prior to incubation. J. Exp. Zool. 93: 51.

LAWRENCE, A. S. C., M. MIALL, J. NEEDHAM, & S. C. SHEN. 1944. Studies on the anomalous viscosity and flow-birefringence of protein solutions. J. Gen. Physiol. 27: 233.

LEATHES, J. B. 1925. Role of fats in vital phenomena. Lancet 853, 957.

LEWIS, W. H. 1933. Locomotion of rat lymphocytes in tissue cultures. Bull. Johns Hopkins Hosp. 53: 147.

1942. The Relation of the Viscosity Changes of Protoplasm to Amoeboid Locomotion and Cell Division. In: *The Structure of Protoplasm.* Iowa State College Press.

LOEB, L. 1928. Amoeboid tissue and amoeboid movement. Protopl. 4: 596.

MAST, S. O. 1926. Structure, movement, locomotion and stimulation in *Amoeba.* J. Morphol. & Physiol. 41: 342.

1941. Motor Response in Unicellular Animals. In: *Protozoa in Biological Research.* Columbia Univ. Press. New York.

MONNÉ, L. 1941a. Über Farben und Farbveränderungen lebender Zellen im Dunkelfeld. Protopl. 36: 222.

1941b. Polarisationsoptische Analyse des Zytoplasmas der Spermatozyten von *Lithobius forficatus* L. Ark. Zool. 34B: 1.

1944. Cytoplasmic structure and cleavage pattern of the sea urchin egg. Ark. Zool. 35A (13): 100.

1946. Struktur und Funktionszusammenhang des Zytoplasmas. Experientia 2: 1.

MONROY. A. 1947. Further observations on the fine structure of the cortical layer of unfertilized and fertilized sea urchin eggs. J. Cell. & Comp. Physiol. 30: 105.

MONROY, A., & A. MONROY ODDO. 1946. Richerche sulla fisiologia della fecondazione. I. Pubbl. Staz. Zool. Napoli 20: 46.

MOORE, A. R. 1928. On the hyaline membrane and hyaline droplets of the fertilized egg of the sea urchin, *Strongylocentrotus purpuratus.* Protopl. 3: 524.

MORGAN, T. H. 1903. The relation between normal and abnormal development of the embryo of the frog, as determined by the effect of lithium chloride in solution. Arch. Entw.-Mech. 16: 691.

MOTOMURA, J. 1933. On the presence of the immovable cortical cytoplasm in the centrifuged sea urchin egg and its importance on the determination of the polarity. Tohoku Imp. Univ. Biol. 8.

1935. Determination of the embryonic axis in the eggs of Amphibia and Echinoderms. Sci. Rep. Tohoku Univ. (5. Ser., Biol.) 10. Sandai, Japan.

NEEDHAM, J. 1942. *Biochemistry and Morphogenesis.* Cambridge Univ. Press.

PALMER, K. J., & F. O. SCHMITT. 1941. X-ray diffraction studies of lipide emulsions. J. Cell. & Comp. Physiol. 17: 385.

PALMER, K. J., F. O. SCHMITT, & E. CHARGAFF. 1941. X-ray diffraction studies of certain lipide-protein complexes. J. Cell. & Comp. Physiol. 18: 43.

PANTIN, C. F. A. 1923. On the physiology of amoeboid movement. J. Marine Biol. Assoc. 13: 24.

1926. On the physiology of amoeboid movement. Brit. J. Exp. Biol. 3: 275.

PARPART, A. K., & A. J. DZIEMIAN. 1940. The chemical composition of the red cell membrane. Symp. Quant. Biol. 8: 17.

PICKEN, L. E. R. 1940. The fine structure of biological systems. Biol. Rev. 15: 133.

Rhumbler, L. 1898. Physikalische Analyse der Lebenserscheinungen der Zelle. Arch. Entw.-Mech. 7: 103.

Runnström, J. 1928. Die Veränderungen der Plasmakolloide bei der Entwicklungserregung des Seeigeleies. Protopl. 4: 388.

Runnstrom, J., & L. Monné. 1945. On some properties of the surface layers of immature and mature sea urchin eggs, etc. Ark. Zool. 36A (18): 1.

1945. On some changes in the properties of the surface layers of the sea urchin egg due to varying external conditions. Ark. Zool. 36A (20): 1.

Runnström, J., L. Monné, & L. Broman. 1943. On some properties of the surface layers in the sea urchin egg and their changes upon activation. Ark. Zool. 35A (3): 100.

Scarth, G. W. 1927. The structural organization of plant protoplasm in the light of micrurgy. Protopl. 2: 189.

Scarth, G. W., J. Levitt, & D. Siminovich. 1940. Plasma membrane structure in the light of frost-hardening changes. Symp. Quant. Biol. 8: 102.

Schaeffer, H. A. 1920. *Amoeboid Movement.* Princeton University Press.

Schechtman, A. M. 1937. Localized cortical growth as the immediate cause of cell division. Science 85: 222.

Schmidt, W. J. 1937. Die Doppelbrechung von Karyoplasma, Zytoplasma und Metaplasma. Protopl. Monogr. Borntraeger, Berlin.

Schmitt, F. O. 1941. Some protein patterns in cells. 3. Growth Symp. : 1.

1944. Structural Proteins in Cells and Tissues. In: *Advances in Protein Chemistry* 1: 27.

Schmitt, F. O., & R. S. Bear. 1939. The ultrastructure of the nerve axon sheath. Biol. Rev. 14: 27.

Schmitt, F. O., R. S. Bear, & E. Ponder. 1936. Optical properties of the red cell membrane. J. Cell. & Comp. Physiol. 9: 89.

1938. The red cell envelope considered as a Wiener mixed body. J. Cell. & Comp. Physiol. 11: 309.

Schmitt, F. O., C. E. Hall, & M. A. Jakus. 1943. The ultrastructure of protoplasmic fibrils. Biol. Symp. 10: 261.

Schulman, J. H., & E. K. Rideal. 1937. The action of haemolytic and agglutinating agents on lipo-protein monolayers. Proc. Roy. Soc. London B 122: 46.

1937. Molecular interaction in monolayers. Proc. Roy. Soc. London B 122: 29.

Seifriz, W. 1927. Protoplasmic papillae of *Echinarachnius* oocytes. Protopl. 1: 1.

Spek, J. 1918a. Die amöboiden Bewegungen and Strömungen in den Eizellen einiger Nematoden während der Vereinigung der Vorkerne. Arch. Entw-Mech. 44: 217.

1918b. Oberflächenspannungsdifferenzen als eine Ursache der Zellteilung. Arch. Entw.-Mech. 44: 1.

Waugh, D. F., & F. O. Schmitt. 1940. Investigations of the thickness and ultrastructure of cellular membranes by the analytical leptoscope. Symp. Quant. Biol. 8: 233.

Weiss, P. 1941. Nerve patterns: the mechanics of nerve growth. Third Growth Symp. Growth 5 (Suppl.): 163.

1947. The problem of specificity in growth and development. Yale J. Biol. & Med. 19: 235.

Wilber, C. G. 1946. Notes on locomotion in *Pelomyxa carolinensis.* Trans. Am. Micr. Soc. 66: 319.

Witschi, E. 1930. Experimentally produced neoplasms in the frog. Proc. Soc. Exp. Biol. & Med. 27: 475.

Yamamoto, T. O. 1940. Ryhthmical contractile movement of eggs of trouts. Annot. Zool. Jap. 19: 69.

Discussion of the Paper

DR. ROBERT CHAMBERS (*New York University, New York, N. Y.*):

The study of imitations of living processes by the use of non-living material has its good points but requires caution. The plasma-membrane is a structure of which we still know practically nothing. To compare it with a lecithin-like film or even to regard it as a lipoprotein is still a questionable procedure. The plasma-membrane may extraordinarily resemble a lecithin film, but its reaction to electrolytes is quite different. Dr. Holtfreter pointed out that the presence of calcium in the medium stiffens a lecithin film and suggested that this offers still another resemblance to the cell membrane. Let us consider the sea urchin egg. The external component of the surface of the egg can be shown to be an extraneous coat which overlies the "plasma-membrane" and which is removable without detriment to the life of the cell. On the other hand, the selectively permeable and physiologically essential component of the surface is fluid in the presence of calcium and can be shown to be highly liquid when the calcium in the medium is in excess. It is the extraneous coat and not the protoplasmic surface film or the so-called plasma-membrane which is stiffened in the presence of calcium.

DR. DOUGLAS MARSLAND (*New York University, New York, N. Y.*):

I was most interested in Dr. Holtfreter's observations bearing on the physiology of amoeboid movement. However, I find it difficult to agree with his conclusion that the force of this movement originates in the membrane of the cell rather than in the plasmagel layer.

One difficulty lies in the fact that the total tension developed in the cell membrane—as measured in a variety of cells, by Harvey, Cole and others—is of a very low order, not in any case exceeding about 5 dynes per centimeter; and such a force does not seem adequate to motivate the movements of the cell. Radical deformation of the shape of the cell by so small a force would not be possible, especially at such times as the plasmagel layer is firmly set—as it is while active movements are progressing. Also, Dr. Holtfreter implies that the force that leads to cleavage likewise resides in the membrane; yet, at the time when the cleavage furrow cuts through the cell, the cortical plasmagel layer of the egg displays a maximum firmness which resists any displacement of cortical granules even by relatively high centrifugal forces.

A second difficulty is that solation and gelation are not visible processes which can be observed directly under the microscope. The hyaline protoplasm immediately subjacent to the cell membrane is capable of undergoing gelation and of developing contractile forces, although Dr. Holtfreter seems to assume that this hyaline layer is always in the state of a sol.

The cell membrane is important, no doubt, as regards the orientation of amoeboid movement, since any local diminution of the membrane tension would be conducive to the outbreak of a pseudopodium at the weakened point. Also, the firmness of attachment between the cell membrane and the subjacent plasmagel must be important, since a pseudopodium can form only in areas where detachment occurs. Such a detachment permits an outflow of plasmasol into the incipient pseudopodium. If a residual layer of plasmagel persists subjacent to the area of detachment, the outflowing sol is sieved free of granules and is perfectly hyaline. However, sometimes the "plasmagel sieve" disintegrates entirely, and then the outflowing sol is filled with granules. But whether it is hyaline or granular, the plasmasol of the pseudopodium has a capacity to gelate, and this gelation guides and supports the pseudopodium and finally limits the extension and permits the pseudopodium to undergo retraction.

The dissociated blastomeres of the amphibian embryo provide an interesting material for the study of amoeboid movement. Free-living amoebae sometimes display a similar set of unusual patterns of movement, but not as frequently or as plainly. However, to discard the sol-gel interpretation of amoeboid movement on the basis of these new observations does not seem either necessary or justifiable.

DR. WARREN H. LEWIS (*The Wistar Institute of Anatomy and Biology, Philadelphia, Pa.*):

Little has been said, so far, about the mechanics of development or the forces involved in morphogenesis. Dr. Holtfreter has considered the role of the "cell membrane" in the mechanics of cell movements. I do not agree with some of his ideas, but there is not time to consider them here, as I wish to deal especially with the mechanics of development and morphogenesis which have not been mentioned.

All eggs and cells have superficial gel layers (gel layer for short) which exert continuous contractile tension, a fundamental property of protoplasm when in the gel state. Gel layer and endoplasm are reversible states of the cytoplasm. The superficial gel layer corresponds more or less to the cell membrane of Holtfreter and to the plasmogel plus plasmolemma of the amoeba. It consists of a highly viscous surface layer that shades into somewhat less viscous gel layer. The surface layer corresponds to plasmolemma of Mast and Holtfreter's "surface coat" of amphibian eggs. It is part of the superficial gel layer and might be designated as surface coat. The superficial gel layer plays a leading role in the mechanics of development. Its contractile energy is one of the principal forces involved in morphogenesis.

Local increases and decreases of the contractile tension of the gel layer are responsible for changes of cell form, cell locomotion, flow of endoplasm, cleavage of cells, and (during development) for the cleavage of the egg into its many cells—every cell of the blastula and of even later stages has, for a time, some of the superficial gel layer of the one-celled egg carried into the depths by the contraction of an equatorial band of the gel layer which produces the cleavage furrow. They also account for infiltration and interpenetration by migration of individual cells among others of their own and of different types and for other equally important processes.

When firmly adherent cells of epithelial membranes act together, quite different phenomena result. An increase of the contractile tension on one surface of an area of adherent cells that resist distortion will result in a concave depression, or invagination, on the side of the greater tension. Invagination plays an essential role in early development of blastopore, neural tube, optic vesicle, optic cup, lens, otic vesicle, nasal pit, and probably all other organs that arise from epithelial membranes and tubes.

In addition to cell migrations and invaginations, pulls, pushes, and squeezes produced by contractions of gel layers play essential roles in development.

The zebra fish egg has a strong gel layer that exerts contractile tension in all tangential directions. Contraction of the yolk part squeezes endoplasm out of the yolk to form the disk as its part of the gel layer relaxes. Large yolk globules are held back by a sieve zone. After endoplasm is squeezed out of the yolk, its globules are compressed into polyhedrons by continued contraction of yolk gel layer. During gastrulation, contraction of the yolk part of the gel layer pulls the attached blastodisk over the yolk and at the same time thrusts (pushes) the latter against the disk. As yolk gel layer contracts, it solates and by the time disk edge is pulled to the vegetal pole it has entirely solated. Inner active disk cells migrate around its edge to produce involuting germ ring.

All surface cells of *Amblystoma* blastulae are strongly adherent to one another by their superficial gel layers. The presumptive areas, vegetal pole, dorsal entoderm, ventral entoderm, mesoblast, chordablast, neuroblast, and ectoblast behave differently.

Most vegetal pole area cells migrate inward away from the surface. Some invaginate with adjacent dorsal entoderm at the blastopore and produce a shallow archenteron. As the superficial ends of vegetal pole and dorsal entoderm cells contract, they pull adjacent adherent ventral entoderm toward vegetal pole and blastopore and chordablast to the edge of the dorsal lip. These, in turn, pull mesoblast and ectoneuroblast toward the vegetal pole and blastopore.

The arched dorsal lip contracts and advances and pulls row after row of chordablast cells over its edge against ventral entoderm, like the treads of a caterpillar tractor turning to the ground. This elongates chordablast, extends archenteron posteriorly, covers endoderm (yolk plug), and reduces blastopore size. As chordablast is turned under, it pulls adherent neuroblast toward and finally to the blastopore edge. Neuroblast is elongated and narrowed by this pull.

In the meantime, the superficial ends of the ventral entoderm cells contract and reduce its area. This pulls adherent mesoblast and ectoblast toward the blastopore. The superficial ends of mesoblast cells at the mesoentoderm line contract and the cells migrate inward. This pulls neighboring mesoblast cells to this line where they, in turn, migrate inward until all leave the surface. This pulls adherent ectoblast to the blastopore. Mesoderm ingression and dorsal lip contraction result in a contrasting constriction around the mesoderm-dorsal lip circle which is partly responsible for dorsal lip advance and blastopore closure.

These vegetal hemisphere contractions reduce its area and push or thrust yolk mass against blastocoele and ectoneuroblast and help to stretch the latter as it is pulled toward the blastopore.

Archenteron expansion is probably produced by pressure of fluid secreted into it by its lining cells. Secretion of fluid plays an active mechanical part in the development of blastocoele, mammalian blastocyst, brain ventricles, optic cup, ear vesicle, etc.

The factors responsible for the superficial gelation of the cytoplasm and for local changes of its contractile tension are unknown.

DR. J. HOLTFRETER:

In reply to Dr. Chambers, I should like to refer to what I said with regard to the difficulty of determining by microscopic observation, whether or not such a delicate structure as the plasma membrane has the properties of a liquid (see footnote, p. 739). Apparently, all workers who have tried to elucidate the physico-chemical properties of this structure agree that the outermost layer of living cells is organized in such a way that its molecular constituents cannot move at random, as in a liquid phase. If the concept of a paracrystalline lipoprotein nature of the plasma membrane be accepted, then churning or other movements observed at the surface of denuded eggs should be due to variations in the molecular packing of the organized surface layer. Furthermore, analogous to the translations occurring within the wall of a myelin body, "flowing" motions in the cell surface might be caused by the slipping of a molecular surface lamella over a subjacent lamella. As far as the embryonic amphibian cells are concerned, their living outer membrane appears to have normally a semi-solid consistency. Without the support of the endoplasmic gel wall, it can spontaneously exhibit folds of elastic tension, mobile ruffles, and long tubular or attenuated pseudopodia. This seems to indicate that the membrane is not fluid, and that other forces than surface tension determine its shape and motility. However, prolonged exposure to salt solutions lacking calcium renders the membrane flabby, non-contractile, and highly permeable, whereas, in solutions of a sufficiently high calcium concentration, the membrane shrinks into a definitely solid substance which resists being cut with a glass needle.

With reference to the remarks of Dr. Marsland, I should like to point out that the extraordinary size of embryonic amphibian cells makes it easy to observe their protoplasmic architecture under the microscope. In fact, both the structural and kinetic details in these cells are so spectacular that my undergraduates in embryology have no difficulty in seeing them. Using the well-recognized criterion of the presence or absence of unrestricted Brownian movement, one can readily decide whether the granulated endoplasm is in a sol or gel state. I agree, however, with Dr. Marsland that in more highly differentiated cells, when granulation diminishes and the plasmagel becomes more firmly associated with the cell membrane, these two layers can no longer be optically differentiated, except at the tip of a pseudopod, where they remain separated by a hyaline layer referred to in my paper as "ectoplasmic fluid." Since both this fluid and the plasmagel may be hyaline, their optical appearance is, of course, no criterion for the physical state of the protoplasm.

In the comparatively minute cells isolated from mammals, even such an experienced microscopist as Dr. W. H. Lewis (1942) has not always been able to recognize a gelated cortex, nor could he detect a reversible sol-gel formation when these cells were migrating. Such material is, therefore, unfavorable for an analysis of amoeboid movement. On the other hand, Dr. Lewis has published beautiful microphotographs of the advancing tip of a slime mold (1942, Figures 1–4) which show clearly that in this organism, as in *Amoeba* and amphibian cells, the frontal wall of the plasmagel is separated from the outer membrane by a spacious layer of a hyaline substance, said to be fluid. Furthermore, the outer membrane can be seen to perform lively amoeboid movements while the subjacent plasmagel wall shows a smooth surface, a fact which obviously does not fit into the sol-gel interpretation held by Dr. Lewis and others.

The various patterns of movement occurring in isolated amphibian cells of different age and tissue derivation are not unusual, since they comprise most types of amoeboid movement observed in protozoa and somatic cells. Whether the latter can be interpreted along the lines here suggested remains to be investigated. At any rate, the current sol-gel theory fails consistently to explain the various kinetic phenomena here described, whereas all these phenomena are readily covered by the membrane theory advanced as an alternative. Apparently, the mechanism here proposed is likewise applicable to the interpretation of ciliary movements.

Lack of space prevents me from discussing the ideas of Dr. Lewis on morphogenetic movements. To avoid misunderstandings, I only wish to restate my conclusion that the protoplasmic structures defined respectively as coat, cell membrane (plasmalemma), and plasmagel, differ from each other structurally as well as functionally. In cases where these layers cannot be distinguished from each other, one should not ascribe arbitrarily any function to any one of them which they probably do not perform. I am not aware of experimental data showing that the coat, or any other superficial "gel layer," has essentially more to do with the mechanics of gastrulation than has been mentioned in my foregoing account.

PROSPECTIVE AREAS AND DIFFERENTIATION POTENCIES IN THE CHICK BLASTODERM

By DOROTHEA RUDNICK
Osborn Zoological Laboratory, Yale University, and Albertus Magnus College, New Haven, Connecticut

EMBRYOLOGISTS interested in the chick, like everyone else in the field, find an expanding technology opening up before them. In the last decade, even during the war, microchemical, histochemical, and cytochemical tools have emerged and have developed to a degree that makes it immediately possible to obtain accurate knowledge of even the earliest events in development. It is imperative for embryologists to make certain of the definition of their morphogenetic problems, in order that the new tools be applied precisely and profitably. The present opportunity to examine the factual basis of some classical concepts as applied to the chick is thus very welcome.

Germinal Movements during Gastrulation. The discussion will center around problems of the nervous system, commencing with the relationships of the medullary area to the underlying axial mesoderm. It is characteristic of our uncertainty of the course of early embryonic movements in the chick that Dr. Spratt of Johns Hopkins should have been able, just recently, to make us readjust all our ideas of early localization by his marking experiments on the unincubated blastoderm.

It will be well to recall that gastrulation in the chick is performed in at least two steps, perhaps three. The first, which has progressed to a variable degree at the time of laying, is the delamination of the lower layer or hypoblast, from an originally single blastodermal sheet. This step takes place without much superficial morphological indication. The hypoblast customarily is believed to give rise to the entoderm. The mesoderm is invaginated from the upper layer mainly by migration through a morphologically patent blastopore, the primitive streak, which takes shape in correlation with other visible changes in the embryonic area. Descriptions are available of some mesoderm, in early stages, rather delaminating irregularly from the upper layer than invaginating in orderly fashion through the streak. The first appearance of the streak itself presents this picture, which might be considered a preliminary or accessory manner of mesoderm formation.

In view of the current interest in the association of carbohydrate metabolism with invagination, it is surprising that Jacobson (1938) has thus far been the only investigator to describe the invagination or delamination process in the chick as being accompanied by loss of glycogen in the cells involved. This would make the situation comparable to that in the amphibian gastrula, and thus far seems to indicate the only qualitative chemical difference that can be correlated with mesoderm formation.

The primitive streak arises by movement of the epiblast layer to a median posterior line; material is piled up on either side of this line; eventually, a mesodermal sheet grows out underneath, spreading at first laterally and posteriorly. Only after the streak has reached its maximum size does an appreciable anterior migration of continuous mesoderm take place, *i.e.*, the head-process with its paraxial mesoderm. This sheet is preceded by some looser, less coherent mesoderm, the "anterior free mesoderm" or, essentially, the prechordal plate, which has been briefly described by Pasteels (1936) as invaginating also through the anterior streak, ahead of the chorda.

Spratt's (1946) recent study shows with great exactness the movement of carbon marks placed on the surface of the unincubated blastoderm cultivated *in vitro*. His Figure 13 summarizes the information in a coherent diagram: it is clear that the whole posterior half of the unincubated blastoderm eventually becomes mesoderm or streak, whereas the posterior border of the anterior half is the site of a graded but—in peripheral regions—very marked stretching or expansion, considered to be an active rather than a passive process. It would seem that we have at last a definitive answer to the question of how much epiblast material goes into the streak—a question that has received widely differing answers, not only from nineteenth-century workers but from recent experimental workers as well (Wetzel, 1929; Pasteels, 1936). The fact that the work of Spratt was done *in vitro* may perhaps invite objection, but the author controlled his observations by careful measurements, comparing his preparations with normally growing ones, and by many repetitions. If, working in this artificial environment, he had found a less energetic or less extensive invagination than had previously been described, some reservation of judgment might perhaps be justified, since explantation notoriously retards or inhibits many processes. As the case is, however, one feels sure that the displacements are at least as radical as Spratt indicates.

Moog (1944) briefly describes gradients in phosphatase (especially in acid phosphatase) in early blastoderms that might possibly coincide with this concentration of material in the streak. It would be of particular interest to follow the history of regions that evidently expand most actively during this period, to determine if they are the site of enzymatic loss, as might be suggested by her findings.

The disposition of areas of the future embryo in the stage of the maximum or definitive primitive streak is well known from the maps of Wetzel and Pasteels. This plan may be re-projected back on the unincubated blastoderm, with results somewhat as in FIGURE 1. In this figure, the right half shows merely the boundaries of the streak at its inception, and its definitive stage when the primitive pit, surrounded by Hensen's Node, occupies a position immediately behind the stationary midpoint of the blastoderm. The limit of invagination in post-definitive streak stages has also been estimated from Spratt's 1947 figures. The left side of the drawing attempts to show the organ-forming areas, on the basis of this information. The medullary area must lie anterior to the invaginating region, and the mesodermal areas must enter the streak in inverse order as compared to their final position. There is no accurate experimental basis for boundaries between the mesodermal areas, except possibly in the case of the chorda.

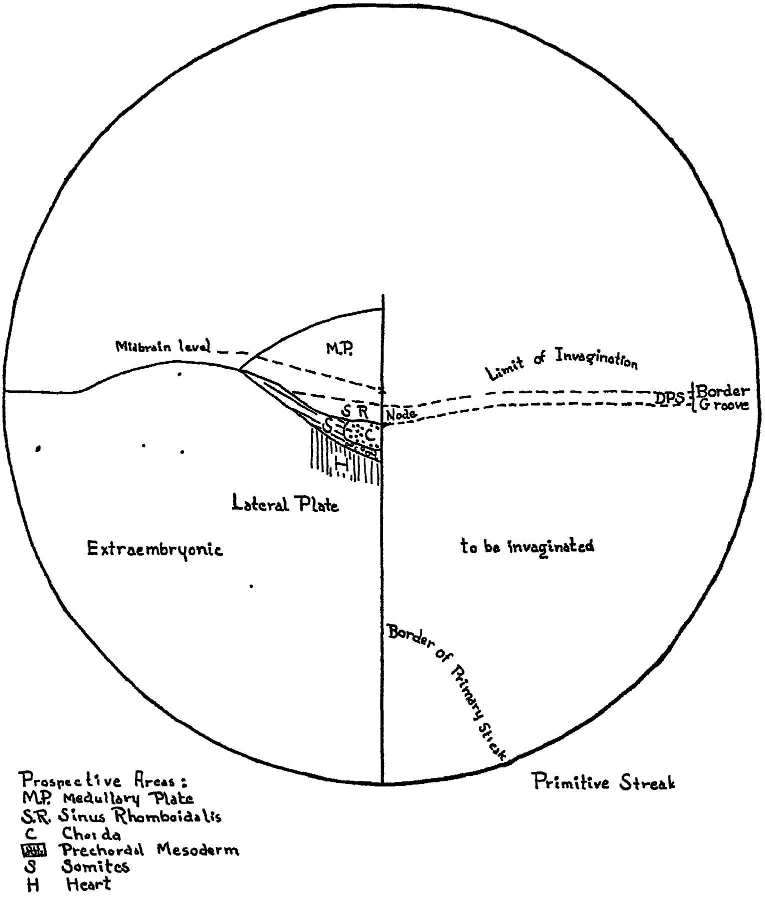


FIGURE 1. Diagram of prospective areas in the unincubated blastoderm, in the light of recent experimental results. Right side shows the primitive streak material, and limits of invagination. Left shows prospective embryonic areas. Intramesodermal boundaries are all hypothetical.

The line between the mesodermal area and the medullary region has a clear basis in Spratt's records. During formation of the streak, the shape of the medullary area must change from something like a half-circle to a long ovoid, largely through stretching and wheeling of the latero-posterior borders. As the head-process grows, the anterior half of the oval will elongate through differential growth greatest in the midline. This process moves Hensen's Node and the primitive streak relatively posterior, and swings segment after segment into rectilinear transverse position. At the stage of the definitive primitive streak, all of the medullary plate that is really laid out, as a flat area, consists of the anterior and lateral parts. The median material, or neural floor, is still to be proliferated from the anterior border of Hensen's Node, along with the notochord.

Thus, different parts of the prospective medullary tube have definitely different histories. Some are older than others, as epithelial areas. The significance of this is evident from the consideration that the impetus toward differentiation of the medullary plate is, in all probability, given by the emigrating mesoderm that comes to underlie the ectodermal layer. The pattern of emigration of mesoderm from the streak is not known as precisely as is the history of the superficial layer, but the general lines can be sketched from morphological reasoning and from the vital stain experiments of Pasteels. The most anterior part of the medullary area is never underlain by more than diffuse prechordal material. The lateral and posterior parts, on the contrary, are underlain by a succession of mesodermal areas, as the sheet migrates laterally: first, probably, extra-embryonic blood-forming mesoderm; then heart and lateral plate; only later parachordal or somite mesoderm takes its definitive station. We definitely do not know the rate or exact directions of this migration, but its occurrence is beyond doubt. As for the neural floor, it seems to be evolved from the front of Hensen's Node *pari passu* with the underlying chorda. The two adhere closely.

Induction of Medullary Plate. The next consideration is a brief review of the evidence, for the chick, that medullary plate arises as a result of induction from an underlying layer. It will be recalled that the original proof of the dependence of the amphibian medullary plate on a stimulus from the underlying archenteron roof consisted broadly of two types of experiment: (1) the test of the ability of carefully specified germinal regions (blastopore lip, archenteron roof or later derivatives) to induce medullary plate in ectoderm that would normally differentiate into something else; and (2) the test of the differentiation capacity of isolated ectoderm, before and after it is underlain by inducing mesoderm. The second test was carried out by various isolation or transplantation methods, culminating in the systematic area-by-area investigation reported by Holtfreter in 1936.

For the chick, Waddington (1930 *et seq.*) has shown that the primitive streak and, to some extent, its derivatives have the capacity to induce supernumerary medullary plate in ectoderm of widely varying regions of the blastoderm. Woodside (1937) has made clear at what age it is possible to elicit this reaction, *viz.*, response is maximal if the graft has an opportunity to act while the host streak is developing, and virtually disappears when the host is forming neural folds, lingering longer in anterior regions of the blastoderm than in posterior. It is not easy to verify, from published accounts and figures, if every part of the non-medullary ectoderm has been so tested and has proved capable of forming medullary plate when properly stimulated. Certainly, lateral parts of the *area pellucida* ectoderm and some—presumably all—of that of the *area opaca* can so respond. There seems to have been no critical test of the *area pellucida* ectoderm just anterior to the normal axis, a region peculiar in that, normally, it is never underlain by mesoderm. This test would be of interest. Nevertheless, it seems safe to say that all ectoderm of suitable age can be induced to form medullary plate and to undergo some regional differentiation.

As to the efficient stimulus, the very early streak seems not to have been tested for inducing capacity; but different parts of later streaks, and of the head-process with overlying ectoderm, have proved effective. The complexity of the inductions produced by these non-homogenous transplants acting in the growth field of a host embryo *in vitro* is extreme. This is perhaps clearest in the long series of cases reported by Waddington and Schmidt (1933), which were also heteroplastic (duck to chick and *vice versa*). In this report, the cases in which unmistakable regional character or even polarity can be ascertained in the induced medullary plate are relatively few. From these few cases, however, the authors have shown some evidence for differences in character of the induction, depending on the region (anterior or mid-streak, head-process, etc.) from which the graft was taken.

There is, thus, satisfactory evidence that medullary induction can occur in the chick ectoderm. We may proceed to the evidence that this process occurs normally. For the chick, in the stages when ectoderm is being underlain by mesoderm, separation of the two layers for experimental tests has not been technically possible. The critical, direct test of differentiating capacity of the same region of ectoderm, before and shortly after mesodermal contact, has not been performed. Only indirect approaches can be cited and they have led to new questions rather than to satisfactory answers.

Some time ago (Rudnick, 1938a), separate pieces of the early streak blastoderm were grown *in vitro*. In the earliest series, cuts were calculated—by the information then available but quite coinciding with Spratt's new data—to separate the forming streak, visibly invaginating, from superficial mesoderm about to be invaginated, and these in turn from ectoderm. No embryonic axis differentiated from any piece, but merely little clumps or vesicles of tissue. Few tissues could form after this drastic partition, namely, generalized tissues like blood, mesenchyme, epithelium, and, in addition, heart muscle and medullary plate. Medullary plate, or rather little tubes or spheres, developed from the most anterior (ectodermal) explants in a certain percentage of cases. It is not certain whether they did so from other pieces. At this stage, before the groove of the streak had appeared, it seems impossible that mesoderm invaginated through the streak region could have reached the anterior piece, which, however, contained all or most of the prospective medullary material. In face of this, one is reduced to two radically different alternative hypotheses: either the prospective medullary region, *in toto* or in part, contains potencies for medullary differentiation antedating mesodermal induction; or, in pieces cut and explanted, there is some irregular invagination or substitute therefor, and this process is adequate for induction. Further along this second line of thinking, there is the question of the normal prechordal mesoderm—whether perhaps its regular route of invagination is not by irregular delamination, which could have occurred in the isolated explants; in which case the medullary tubes would represent forebrain.

In this material, there was no indication as to what region or level of the axis the medullary nodules might represent. Even in explants grown from later streaks, from regions where induction by emigrating mesoderm was possible, no regional differences could be ascertained. Clearly such fragments present drastic mechanical opposition to regional morphogenesis.

Spratt (1940, 1942), a short time later, reported experiments in which regional organization had been given a much better chance. These consisted in dividing early blastoderms (pre-streak through head-process or later) into only two parts by a transverse cut. If the cut was made approximately through the midpoint of the blastoderm, in pre-streak or early streak stages, it is obvious that all or part of the prospective medullary plate would be isolated from the intact mesoderm-streak region behind. From cultivation of both pieces, Spratt obtained typically a fairly good axis from the posterior piece in which the streak formed, and a little nodule of medullary plate in the center of the anterior piece. At the time, Spratt interpreted the rather shapeless anterior bits as forebrain, and, in a few cases, on subculturing, obtained retinal pigment from them; the posterior axes he thought specifically deficient in forebrain. One wonders if, in view of his recent localization experiments, he might not wish to revise this view and think of the posterior axes as being inhibited but not necessarily lacking in any one level. It is clear that not only the prospective forebrain but a good bit besides was included in the anterior piece, whereas very little destined medullary material was present in the posterior piece. In development *in vitro*, the anterior cut border of the posterior piece (where lay most of the prospective medullary material that was included) remained fixed to the clot. This emphasizes the conclusion that the medullary plate of the posterior axes must have formed largely from epithelium that normally would have entered the streak to become mesoderm. It would seem that Spratt has shown here, among other things, that prospective mesoderm also can be induced to form medullary plate.

The anterior pieces again present the problem of whether and by what the medullary nodules were induced. To illustrate the difference in result, when essentially the same transection experiment is performed after the whole brain-forming region is underlain by mesoderm, one should consult the figures in Spratt, 1940 (Plate 2). In these cases, the anterior piece clearly complements the posterior, and the mosaic character of the separation of forebrain or its parts from the rest of the axis is obvious.

Thus, attempts to test parts of the medullary field before formation of the mesoderm have served to demonstrate our ignorance of the normal course of invagination, not to mention our ignorance of the organizing substance itself. If, at first glance, things look as if the early medullary field in the chick (the site of the later forebrain) possesses some powers of self-differentiation, it is not likely that such an explanation will satisfy investigators or be allowed to remain without further analysis.

Regional Organization of the Medullary Plate. The forebrain problem raised in the preceding section may serve to introduce some discussion of the pattern present in the early medullary plate. In searching the literature for cases where a forebrain has been experimentally induced, one finds exactly one instance where an indubitable prosencephalon with optic swellings is figured, as contrasted with vague designations of "head" or "head-fold" which usually look very unlike any part of a normal chick brain. This case is shown in Waddington and Schmidt, 1933, Figure 10, and the beautiful induced head lies beside the normal one at about the same antero-posterior level. The transplant performing the induction was from a chick: the anterior half of a medium (*i.e.*, pre-definitive) primitive streak, including Hensen's Node (whether any material anterior to the node was also included, is not stated), placed ventral side up in the right middle region of the duck host, beneath the epiblast. If anything, the induced brain is farther advanced than the host, that is, optic swellings are clear. The transplant itself has differentiated into neural tube and a little unspecified mesoderm. The graft material immediately touching the induced optic swelling is neural tube, which is actually continuous with the host induction. The authors also note that the most anterior part of the induction is not underlain by any graft at all. It seems questionable as to whether the regional character of this induction is to be attributed, as the authors do, entirely to the character of the graft. Its position with relation to the host embryo is too parallel. If forebrain is induced by graft neural tube, as the morphological relations appear to indicate, one would think that the regional character of the induction must be attributable, at least in part, to host influences. Unless pre-nodal material were included in the graft, the normal inductor for telencephalon was not present. The evidence offered by this case for localization within the inducing system is thus very unsatisfactory.

In the head-process blastoderm, each level of the medullary plate is capable of differentiating into a characteristic brain-vesicle, with accurate histological pattern of cell types and layers, when suitably transplanted, along with underlying mesoderm (Rawles, 1936). This antero-posterior pattern also seems to be present in the definitive primitive streak stage, if allowance is made for the condensed nature of the material. In preceding stages, especially in those where it is known that the definitive relations of axial mesoderm to medullary ectoderm have not been attained, no experimental evidence is available. With our current increase of understanding of the history of movements and relative position of the two layers, a real experimental test of their separate roles in the emergence of the antero-posterior pattern will be possible. The evidence found by Spratt, of potency of at least a cellular order to form retinal pigment in anterior ectodermal pieces isolated from the rest of the axis in pre-streak stages (1942), is a suggestion that the idea of pattern in the ectoderm itself must not be neglected.

Transverse organization of the medullary plate, which changes to dorso-ventral order in the embryo, is a process completed, in the cases studied, later than that of the antero-posterior axis. The cases studied are those of the eye (Clarke, 1936) and the neural crest as indicated by pigment cells (Rawles, 1940; Ris, 1941). In both instances, the histological potency to form these lateral or dorsal structures is retained by the isolated median or ventral strip some time after the formation of medullary folds. The exact movements of material in these cases have not been followed. There is a slight chance, for the eye, that a tenuous strip median in the anterior medullary plate, too narrow to be tested by current methods, may, from the first, be unable to form eye tissue, and that the diencephalic floor arises by enlargement of this strip. The situation looks much more like a gradual separation of a continuous eye field into two laterally placed ones by the actual loss of ability of the medial cells to perform a certain histogenetic task. The case of the neural crest melanophores is clearly of this type. This bilateral or dorso-ventral pattern stabilization seems to be a much slower process than medullary induction *per se*. In the amphibian forebrain, the mesoderm has been shown to play a critical role. This would certainly require some localization of pattern in the mesoderm, even if its expression in the ectoderm emerges only slowly.

A morphological basis for a medio-lateral pattern in the invaginated mesoderm is more immediately clear than it is for an antero-posterior one. It has already been pointed out that the lateral parts of the prospective medullary plate are underlain by a progressively migrating mesodermal sheet and are, thus, subject to contact with a succession of mesodermal areas. In the midline, ectoderm and mesoderm are at first massed, without epithelial relations, in Hensen's Node, and grow forward together from that position. This process has been compared to the growth of a tailbud blastema. It is probable that intermediate gradations between these differing processes occur in intermediate regions. Thus, with reference to time and pattern of contact between the two layers, a medio-lateral differential is ready for analysis by more refined transplantation and by cytochemical methods. Furthermore, at the time these patterns are being realized, excised areas can be regenerated from surrounding tissue—even the whole of Hensen's Node (Waddington, 1930 *et seq.*). This offers another means to study the time required for acquisition of new potencies or for the loss of others. These suggestions are made in order to emphasize that, in spite of some of its drawbacks as experimental material in early stages, the chick embryo has definite morphological features which, once understood, promise to make it especially valuable for study of some general problems of vertebrate pattern.

Differentiation Mechanisms. It would seem convenient to distinguish three types of activity concerned in the differentiation of a nervous system from a medullary plate: (1) secretory, responsible for the maintenance of the various cavities of the central nervous system; (2) mass-movements within the whole epithelium, or large groups of its cells, responsible in the first place for closure and shaping of the neural tube, subsequently for formation of layers, cell columns, nuclei, etc., within definite parts of the medullary tube; and (3) differentiation of individual cells in various directions, such as neurons of several types, supporting elements, ependyma, etc. The first activity has not been the subject of adequate study, though the chick embryo would seem to offer an excellent object in which the nature of the barrier between tissues and cavities could be investigated. It is proposed to consider the second and third processes here, briefly, with reference to possible relations one to another.

The induction reaction itself is, of course, of the nature of an epithelial reaction, first visible in the marked heightening of the cells of the induced region. Subsequent processes are also related to the properties of an epithelium, not to single cells; *e.g.*, the changes in form and polarization responsible for the closure of the neural tube. Hobson's (1941) observation that closure, previous to actual fusion of the medullary folds, can be reversed neatly by dehydrating agents such as a drop of glycerine, calls for further analysis. In later stages, the formation of cell layers, so characteristic of various parts of the brain and neural tube, must also be regarded as a series of mass patterns within an epithelium.

As a contrast, we have the cellular differentiation that has been long and closely studied, whereby single cells apparently undergo the characteristic changes involved in forming neurofibrillae and putting forth nerve processes which then acquire peripheral connections. Barron's (1946) recent interpretative study of differentiation in the motor area of the brachial region of the spinal cord may be cited. He finds that certain individual cells of the neural epithelium of the basal plate start differentiation *in situ*, but that this process is correlated with a peripheral migration to the position which will be occupied by the ventral horn. The first pioneer fibers of the ventral root come from such *primary neuroblasts*. In addition, other cells migrate out from the epithelium in an apparently undifferentiated condition, and later become *secondary neuroblasts* by differentiation in striking proximity to the dendritic processes or cell bodies of the primary neuroblasts. Hamburger and his students (*cf.* Hamburger and Keefe, 1944) have already formulated the idea of a cell-to-cell induction, which Barron supports on morphological grounds, as a result of their experimental studies on altering the peripheral load of the brachial or lumbosacral region of the cord. They have shown that the marked development of the motor columns in these two regions is directly dependent on the presence of a periphery with which the motor fibers leaving the cord may connect, and not on any central connections outside of the few segments comprising the limb area of the cord (Hamburger, 1946). Furthermore, the periphery controls the development of the motor columns not by regulating cell divisions or number of cells within the cord, but by controlling cell differentiation, *i.e.*, the formation of motor neuroblasts in the ventral horns. A cell-to-cell mechanism, whereby a pioneer neuron, once it connects with the periphery, puts out one or more dendrites and simultaneously becomes capable of inducing neighboring undifferentiated cells to follow its lead, until the periphery is loaded, provides a most attractive hypothesis, inviting direct experimental investigation. The only lack is a suggestion of what initiates the process in the first place.

The polarization miscroscope studies of Hobson (1941) form a very interesting beginning of optical studies of the neural epithelium just before and after closure of the tube. The indication is of an orientation of protein micells in the long axes of the cells (that is, radiating out from the center of the tube) and lipoid components perpendicular to the protein. A great share of the birefringence was observed in the various membranes (limiting, cell, and nuclear). This particular type of orientation is by no means peculiar to neural epithelium in the chick embryo. It would, however, be most interesting to make similar observations in slightly later stages, when there is more zonation in the neural wall and the pattern of cellular emigration that has just been discussed is better established. It might, thus, be possible to detect differences in content or structure of the migrating individual cells.

Moog's (1943) studies of the phosphatase distribution at various stages of the developing spinal cord show some striking patterns, some of which, such as the median band of alkaline phosphatase, seem very difficult to understand. The preparations are not figured or described minutely enough to ascertain differences of individual cells, so that it is impossible to say if phosphatase gives any promise of being involved in the differentiation under discussion. The program will evidently be to continue to look for cytochemical and cytophysical differences between individual cells in the neural tube, until the specific difference between migrating and non-migrating individuals becomes apparent, as well as subsequent differences between neuroblasts and undifferentiated cells. A cytochemical cell lineage of the elements and derivatives of the neural epithelium is a highly desirable goal which, when available, will perhaps offer the clue to the mechanism or mechanisms involved.

It is possible that the two major activities in neuron formation—emigration and putting forth nerve processes—are merely steps in one single course of surface changes. Under this view, a series of events would be early initiated in all cells of the neural epithelium, which would progress more rapidly in some cells than in others. In cells changing most rapidly, some double threshold would be crossed and those cells would become primary neuroblasts, simultaneously undergoing internal rearrangement, changing external form, and losing contact with their neighbors to slip outside the epithelium. In cells where the process moves somewhat less rapidly, only an intermediate threshold would be attained. Surface changes would go only so far as to permit these cells to migrate out of the epithelium in a generally rounded state, requiring a further stimulus from outside—an induction—to become neuroblasts; these would be the secondary type. Large groups of cells in which either threshold was reached at the same time would provide a basis for cell layers, columns, etc., as in the differentiating brain epithelium.

It would be well to note that these formal relations within the epithelial structure of the neural tube do not hold any exclusive key to differentiation of neurons. Nerve cells may differentiate in transplanted material that has never had an opportunity to form a tube or more than a most irregular and distorted epithelium (Rudnick, 1938b). It is only necessary to consider the neural crest to realize that no epithelial history at all is requisite to neuron differentiation, however important it may be for the larger pattern of the central nervous system.

If it seems that the foregoing review has been more concerned with exposing our ignorance than with recording progress in the understanding of developmental processes in the chick, it can certainly be replied that the morphological task before us is a most exacting one. We are really faced with ascertaining individual cell differences and cellular relations in material in which, until recently, one cell seemed to look just like another. The fact that all cells in the blastoderm or in the neural tube do not behave alike has led to various hypotheses concerning the mechanisms responsible, *viz.*, progressive internal change; cell contact with chemical or physical effect; even production of substances and their transfer over some distance. The sorting-out of these possibilities inevitably requires more and more detailed knowledge of cellular relations in normal development. The morphological questions we are now trying to answer would have been meaningless in the classical period of descriptive morphology, and, in this sense, it can be submitted that progress is being made.

Literature Cited

Barron, Donald. 1946. Observations on the early differentiation of the motor neuroblasts in the spinal cord of the chick. J. Comp. Neurol. 85: 149–170.

Clarke, L. F. 1936. Regional differences in eye-forming capacity of the early chick blastoderm as studied in chorio-allantoic grafts. Physiol. Zool. 9: 102–128.

Hamburger, V., & E. L. Keefe. 1944. The effects of peripheral factors on the proliferation and differentiation in the spinal cord of chick embryos. J. Exp. Zool. 96: 223–242.

Hamburger, V. 1946. Isolation of the brachial segments of the spinal cord of the chick embryo by means of tantalum foil blocks. J. Exp. Zool. 103: 113–142.

Hobson, L. B. 1941. On the ultrastructure of the neural plate and tube of the early chick embryo, with notes on the effects of dehydration. J. Exp. Zool. 88: 107–134.

Holtfreter, J. 1936. Regionale Induktionen in xenoplastisch zusammengesetzten Explanten. Arch. Entw.-mech. 134: 466–550.

Jacobson, W. 1938. The early development of the avian embryo. I. Endoderm formation. II. Mesoderm formation and the distribution of presumptive embryonic material. J. Morphol. 62: 415–432, 445–488.

Moog, Florence. 1943. The distribution of phosphatase in the spinal cord of chick embryos of one to eight days' incubation. Proc. Nat. Acad. Sci. 29: 176–183.

1944. Localizations of alkaline and acid phosphatases in the early embryogenesis of the chick. Biol. Bull. 86: 51–80.

Pasteels, J. 1936. Analyse des mouvements morphogénétiques de gastrulation chez les oiseaux. Acad. Roy. Belg. Bull. V. 22: 737–752.

Rawles, M. E. 1936. A study in the localization of organ-forming areas in the chick blastoderm of the head-process stage. J. Exp. Zool. 72: 271–315.

1940. The pigment-forming potency of early chick blastoderms. Proc. Nat Acad. Sci. 26: 86–94.

Ris, H. 1941. An experimental study of the origin of melanophores in birds. Physiol. Zool. 14: 48–69.

Rudnick, D. 1938a. Differentiation in culture of pieces of the early chick blastoderm. II. J. Exp. Zool. 79: 399–427.

1938b. Contribution to the problem of neurogenic potency in post-nodal isolates from chick blastoderms. J. Exp. Zool. 78: 369–383.

Spratt, N. T. 1940. An *in vitro* analysis of the organization of the eye-forming area in the early chick blastoderm. J. Exp. Zool. 85: 171–209.

1942. Location of organ-specific regions and their relationship to the development of the primitive streak in the early chick blastoderm. J. Exp. Zool. 89: 69–101.

1946. Formation of the primitive streak in the explanted chick blastoderm marked with carbon particles. J. Exp. Zool. 103: 259–304.

1947. Regression and shortening of the primitive streak in the explanted chick blastoderm. J. Exp. Zool. 104: 69–100.

Waddington, C. H. 1930. Developmental mechanics of chicken and duck embryos. Nature 125: 924.

1932. Experiments on the development of chick and duck embryos cultivated *in vitro*. Phil. Trans. Roy. Soc. B 221: 179–230.

Waddington, C. H., & G. A. Schmidt. 1933. Induction by heteroplastic grafts of the primitive streak of birds. Arch. Entw.-mech. 128: 522–563.

Wetzel, R. 1929. Untersuchungen am Hühnchen. Die Entwicklung des Keims während der ersten beiden Bruttage. Arch. Entw.-mech. 119: 188–321.

Woodside, G. L. 1937. The influence of host age on induction in the chick blastoderm. J. Exp. Zool. 75: 259–282.

BIOCHEMICAL DIFFERENTIATION DURING AMPHIBIAN DEVELOPMENT*

By E. J. BOELL

Osborn Zoological Laboratory, Yale University, New Haven, Connecticut

Studies on the Primary Organizer

BIOCHEMICAL studies of the so-called primary organizer or neural inductor of the amphibian embryo have been directed along two main lines. In the first place, following the discovery that induction of the nervous system could be brought about by agencies other than living embryonic tissues, attempts were made to determine, by various implantation techniques, the nature of the chemical substance or substances which presumably were involved in the process. The literature on this subject is not without interest, but, it must be confessed, our knowledge of the chemistry of the natural inductors has not been greatly enhanced by the numerous studies of this type. Nevertheless, it can hardly be doubted that the interaction between the invaginated chorda mesoderm and the presumptive neural plate, leading to the histogenetic and morphogenetic differentiation of the nervous system, involves the operation of specific chemical substances even though their precise nature and the means through which they exert their influence are unknown. Moreover, it seems increasingly clear, especially in view of Holtfreter's (1944, 1945) recent results, that embryonic tissue contains inducing substances in bound or inactive form and that, as suggested by Needham (1942), a number of physical and chemical agents are effective in bringing about their release.

Respiratory Rates of Dorsal Lip and Ventral Ectoderm. The second line of work has been concerned with an analysis of the metabolic characteristics of inducing and non-inducing parts of the embryo in an attempt to determine whether the developmental events associated with the phenomenon of induction could be correlated in any way with special biochemical processes or properties in particular regions of the embryo. It is perhaps natural that some of the earliest efforts in this direction should concern the study of the overall metabolism of various regions of the gastrula as revealed by their total oxygen uptake. The first investigation was undertaken by Brachet (1935), when he measured the effects of localized cautery on the respiratory metabolism of *Rana temporaria* gastrulae. His results indicated that destruction of the cells in the vicinity of the dorsal lip produced a somewhat greater reduction of oxygen

---

* It is a pleasure to record my thanks to Dorothy Burgess Everett for her invaluable assistance in securing the data represented in the section on "The Development of Enzyme Systems in the Embryo," and for her careful preparation of the graphs in this paper.

consumption than when a corresponding area from the opposite side of the embryo was removed. Measurements of the carbon dioxide production of dorsal lip and ventral ectoderm explants yielded essentially similar results (Brachet, 1935). In these studies, quantitative estimations of the amount of tissue used in the comparative tests were not made, so that the results, although highly suggestive, had to be interpreted cautiously. However, in later work (Brachet, 1936), Kjeldahl nitrogen determinations were made on the explants of *Discoglossus* gastrulae after their carbon dioxide production had been measured, and the results indicated in this case that the rate of carbon dioxide production by pieces of gastrula from the dorsal lip region was unquestionably greater than that of ventral ectoderm.

At the same time, the oxygen consumption of explants of *Triton alpestris* gastrulae was measured by Waddington, Needham, and Brachet (1936) in a modified Gerard-Hartline respirometer. In most cases, the dorsal lip was found to have a greater oxygen consumption than the ventral ectoderm, but when respiration was related to dry weight no difference between dorsal lip and ventral ectoderm was noted.

By an ingenious method, Brachet and Shapiro (1937) measured simultaneously the oxygen consumption of the two halves of the intact gastrula of *Rana sylvatica*. The egg was placed in the center of a capillary, so that the respiration of one hemisphere could be directly compared with that of the other. When the gastrula was oriented so that its dorsal-ventral axis was perpendicular to the long axis of the capillary, no difference in the respiratory rates of the two halves of the egg was observed. However, when the dorsal lip region and ventral ectoderm were directed toward opposite ends of the capillary, the respiration of the hemisphere containing the dorsal lip was some 40 per cent greater than that of the opposite side. This does not necessarily indicate that the dorsal lip has a more intense metabolism than the opposite side of the gastrula, for, as Fischer and Hartwig (1938) have pointed out, it is unlikely that the index drops in the two limbs of the capillary were being influenced by equal amounts of respiring tissue.

Following this, Brachet (1939) further investigated the respiration of explants of gastrulae, using large numbers of pieces in the Meyerhof-Schmitt respirometer. Because of the insensitivity of the instrument, respiration had to be followed for 20 to 30 hours. Brachet's results indicated that the respiration of the dorsal lip region of *Discoglossus* gastrulae was, on the average, some 30 per cent greater than that of the ventral ectoderm. Carbon dioxide production averaged 84 per cent greater in the dorsal lip region. In a second series of experiments, Brachet (1939), using the Brachet-Shapiro apparatus for measurements on the intact gastrula, reported that the average value for the respiration of the dorsal lip-containing half of the gastrula of *Rana temporaria* was not significantly different from that of the half containing the ventral ectoderm. However, carbon dioxide production, measured on the intact gastrula by a modification of the same method, was again seen to be much greater in the half containing the dorsal lip.

Fischer and Hartwig (1938), using as many as 60 explants in Warburg respirometers in experiments lasting from 16 to 23 hours, obtained an average figure for the respiration of the dorsal lip of 2.34 cu.mm. per 10 mg. dry weight per hour. The corresponding value for the ventral ectoderm was 1.83. In experiments of shorter duration, *i.e.*, 6 to 9 hours, the difference was less, the averages being respectively 2.13 and 1.92.

Using the Cartesian diver technique, Boell and Needham (1938, 1939) and Boell, Koch, and Needham (1939) measured the oxygen consumption of explants of dorsal lip and ventral ectoderm of two species of amphibia. In these experiments, made on single explants, the respiratory measurements were confined to a period of 3 hours following isolation of the explants and their insertion into the divers. These investigations showed no significant difference in the average respiratory rates of dorsal lip and ventral ectoderm, and confirmation of this result was obtained with *Rana temporaria* by Needham, Rogers, and Shen (1939). In our work, some variation was noted in the relative rates of respiration of the two kinds of tissue from experiment to experiment. Out of a total of 28 tests in which the tissue fragments whose respiration was being compared were isolated from the same gastrula, 9 showed the rate of oxygen uptake of the dorsal lip to be greater than that of ventral ectoderm. In 16 experiments the opposite was true, and in 3 cases the respiratory rates of the two regions were identical. Brachet (1945, p. 389) has suggested that the variation in our results may have been due to the effects of tissue cytolysis at the air-fluid interface because of the small amount (2 cu.mm.) of Holtfreter's solution used in the divers. However, calculation shows that, even with the largest tissue samples, the volume of saline in which the tissue was supported was about 8 times that of the tissue, and in the majority of the experiments this figure was nearer 20. Examination of the tissues in the divers at the end of the experiments revealed that they remained well-healed and could be removed intact for Kjeldahl determinations. Furthermore, in most cases the respiratory rate was constant during the experimental period. "*Ce serait*," states Brachet, "*un indice d'un état physiologique demeuré normal.*" As will appear below, it is highly probable that the variation in our results was due to the fact that the dorsal lip and ventral ectoderm differed in composition from experiment to experiment.

The experiments which have been mentioned in the discussion thus far are summarized in TABLE 1. They were performed on embryonic material from different species and under various experimental conditions. It is unlikely, however, that such factors are solely responsible for the lack of agreement in the results. Brachet (1939) has stated that, in experiments of short duration, the respiratory rates of dorsal lip and ventral ectoderm are essentially identical, but that in longer experiments the respiratory rate of the dorsal lip is greater. This is due, he believes,

TABLE 1

SUMMARY OF WORK ON THE RESPIRATORY RATES OF DORSAL LIP AND VENTRAL ECTODERM IN THE AMPHIBIAN GASTRULA

Investigator	*Material*	*Rate DL*	*Rate VE*	*DL/VE*
Waddington, Needham, & Brachet	*T. alpestris*^{1}	0.23	0.21	1.10
Brachet & Shapiro	*R. sylvatica*^{2}	85	58	1.47
Fischer & Hartwig	*A. mexicanum*^{3}	2.34	1.83	1.28
Fischer & Hartwig	*A. mexicanum*^{3}	2.13	1.92	1.11
Boell & Needham	*Discoglossus*^{4}	4.80	4.93	0.98
Boell & Needham	*A. mexicanum*^{4}	3.21	3.18	1.01
Boell, Koch, & Needham	*A. mexicanum*^{4}	5.3	4.2	1.26
Brachet	*R. temporaria*^{5}	0.164	0.153	1.07
Brachet	*Discoglossus*^{4}	4.1	3.1	1.31
Needham, Rogers, & Shen	*R. temporaria*^{4}	3.74	3.78	0.99

Respiratory rates (µl. O_2 consumed per hour) are based on the following units of tissue: *1.* milligram dry weight; *2.* gram wet weight; *3.* 10 milligrams dry weight; *4.* milligram nitrogen; *5.* not stated by the author.

to the fact that the dorsal lip tissue differentiates more completely after excision from the embryo than ventral ectoderm and that its respiratory rate therefore increases during the process, whereas that of ventral ectoderm does not. Although this seems a plausible explanation, it is likely that the variations in the results reported above are due even more to fundamental differences in technique. Perhaps the factor of greatest importance has been the failure of various workers to use corresponding areas of dorsal lip and ventral ectoderm for comparison. This is understandable, however, when it is remembered that the various presumptive areas of the gastrula have only statistical boundaries. So-called dorsal lip material varies in composition with the age of the gastrula. In the very youngest stages, it consists exclusively of chorda mesoderm, most of which has not yet been invaginated. In later stages, it is composed of a double layer of tissue, the outer one representing presumptive neural plate and the inner layer comprising chorda mesoderm or archenteron roof. In a given isolate of dorsal lip, the relative amounts of mesoderm and ectoderm will vary, accordingly, with the age of the gastrula from which it is taken as well as with the size of the explant. Furthermore, the term "ventral ectoderm" denotes no well-defined zone in the gastrula, and the respiration of a fragment of ectoderm may therefore be expected to vary somewhat, depending upon its position in the gastrula.

Regional Variation in Respiration of Gastrular Explants. In more recent work, some of these considerations have been taken into account, and comparisons have been made of the respiratory rates of presumptive regions taken from various levels of the gastrula (Boell and Nicholas, 1940; Boell, 1942; Barth, 1939, 1942). The results of these experiments are summarized in TABLE 2 and indicate that differences of considerable magnitude appear in the oxygen consumption of different areas from the same gastrula. There seems to be a gradient of respiratory activity

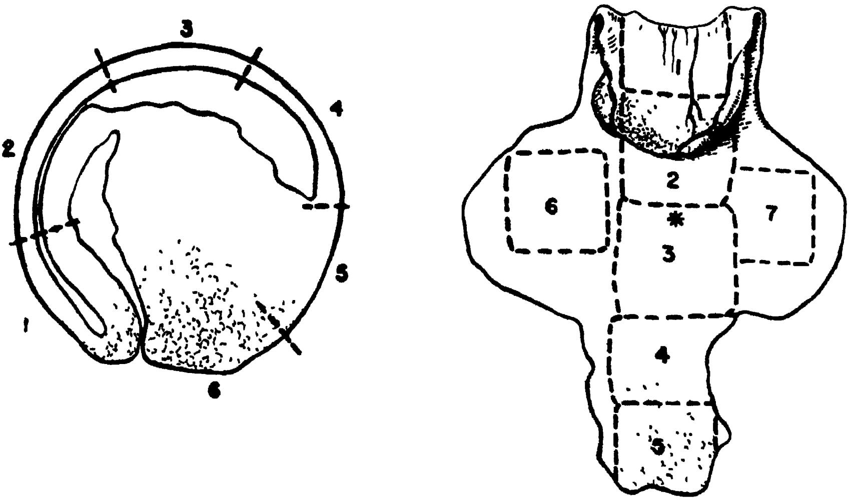


FIGURE 1. Diagram showing method of dissecting pieces of tissue used in the experiments summarized in FIGURE 2 and TABLE 2. Piece 1 refers to the dorsal lip region.

extending from the animal pole to the vegetal pole with the more anterior part of the presumptive neural plate showing the highest rate of respiration. As might be expected, some variation in the absolute levels of the various regions is apparent from gastrula to gastrula, but the general pattern of respiratory rate is similar to that indicated in FIGURE 2.

TABLE 2

RESPIRATION OF VARIOUS REGIONS OF THE GASTRULA

Region	*No. expts.*	*Q'$_{O_2}$*	*P E ave.*
Dorsal lip	31	2.1	0.07
Presumptive neural plate	14	4.9	0.28
Anterior ectoderm	20	4.5	0.16
Posterior ventral ectoderm	10	3.0	0.19
Yolk endoderm	13	1.3	0.09
Right lateral ectoderm	3	3.1	0.11
Left lateral ectoderm	6	3.4	0.31
Chorda mesoderm	6	1.2	0.09

Furthermore, it appears that the differences in respiration among the various regions are more marked when the pieces used for comparison are small and represent well-localized areas. The use of small pieces of tissue in respiration studies has been criticized by Child (1946, p. 127). It may be stated, however, that the pieces of tissue used in the work summarized in TABLE 2 were somewhat larger than those used routinely in transplantation experiments.

Barth (1942) has measured the oxygen uptake of explants of three species of Amphibia by means of a micro-Winkler technique. Although his measurements were made on several gastrular pieces of somewhat larger size than those in our experiments, the results obtained were similar to those already mentioned. It thus appears that the relative respiratory rate of dorsal lip tissue depends upon the particular region

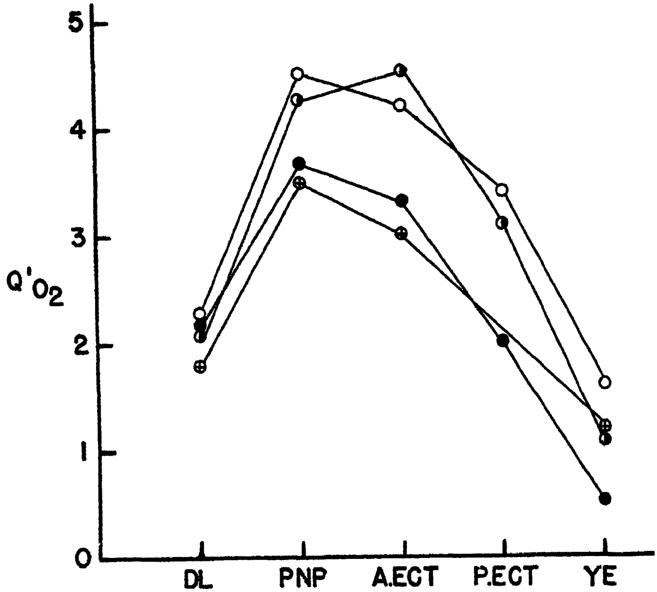


FIGURE 2. Graph showing the variation in the respiratory rates, Q'_{O_2}, of various regions of the gastrula of *Amblystoma punctatum*. DL = dorsal lip material. PNP = presumptive neural plate. A.ECT = ectoderm from the region of the animal pole. P.ECT = ectoderm from region between animal and vegetal pole on ventral side of the gastrula. YE = yolk endoderm.

of the gastrula with which it is compared, being higher than that of ectoderm if the latter tissue is removed from near the vegetal pole, and lower if it is taken from the animal pole of the embryo. This factor is doubtless mainly responsible for the differences in the results of various workers reported in TABLE 1. Moreover, it seems the most likely explanation of the variation in the relative respiratory rates of dorsal lip and ventral ectoderm from gastrula to gastrula reported by Boell and Needham (1939).

Dorsal lip material represents a double layer of tissue while the other regions selected for study are generally thinner. One might question, therefore, whether sufficient oxygen from the air in the gas space of the divers was diffusing into the dorsal lip tissue to maintain respiration at its potential maximum. Therefore, a number of experiments were performed in which the divers were filled with oxygen instead of air and data obtained on the relative respiratory rates of the various regions of the gastrula under these conditions. Comparative figures are presented in TABLE 3. These experiments show clearly that the difference between dorsal lip and ectoderm is due to intrinsic dissimilarities in the tissues rather than to differences in the availability of oxygen. The lower respir-

TABLE 3

COMPARISON OF THE RESPIRATORY RATES OF REGIONS OF THE GASTRULA IN AIR AND IN PURE OXYGEN

Region	*Air*		*Oxygen*	
	Q'_{O_2}	*Per cent*	Q'_{O_2}	*Per cent*
Dorsal lip	2.1	43	1.6	50
P. neural plate	4.9	100	3.2	100
Ant. ectoderm	4.5	92	3.2	100
Post. ectoderm	3.0	57	1.8	56
Yolk endoderm	1.3	27	0.8	25

atory rates in oxygen as compared with the average values in air are simply a reflection of the variation in oxygen uptake in different gastrulae. Pure oxygen, apparently, has no deleterious effect on the tissue explants, and it may be recalled that Parnas and Krasinka (1921) showed that pure oxygen was without effect on the respiration of intact embryos.

Regional Distribution of Yolk and the Respiratory Rate of Yolk-Free "Active" Material in the Gastrula. The amphibian egg contains a large amount of yolk which is distributed asymmetrically between the animal and vegetal poles (Bragg, 1939; Daniel and Yarwood, 1939). Although important as raw material in developmental processes, yolk is generally believed to be relatively inert metabolically. Pickford (1943) found that the dipeptidase activity of yolk is negligible, and Boell and Shen (1944) have shown that yolk obtained by centrifuging homogenates of whole amphibian embryos possessed very little cholinesterase activity. In view of the differences in the yolk content of the various regions of the gastrula, it seemed likely that the differences in respiratory rate might simply reflect the presence of varying amounts of yolk.

Accordingly, determinations of the yolk content of pieces of the amphibian gastrula corresponding to those used in respiration experiments were made (Barth, 1942; Boell, 1942). In the latter work, pieces of gastrula were centrifuged in the Beams air-turbine centrifuge after their initial volume had been determined. Centrifugation at high speeds completely disrupted the tissues and stratified their components so that the amount of yolk could be estimated volumetrically (FIGURE 3). The yolk content of the various regions of the gastrula is summarized in TABLE 4. These figures differ rather widely from the indirect determinations of Barth (1942), but accord with the older measurements of McClendon (cited by Needham, 1931), who found that about 78 per cent of the *Rana pipiens* egg is represented by the yolk fraction.

From the data in TABLES 2 and 4, it appears that the respiratory rate of a piece of tissue is inversely proportional to its yolk content or directly proportional to the so-called "active" material (FIGURE 4). If a direct relationship actually exists between respiration and active substance, one might reasonably expect the extrapolated curve in FIGURE 4 to intersect

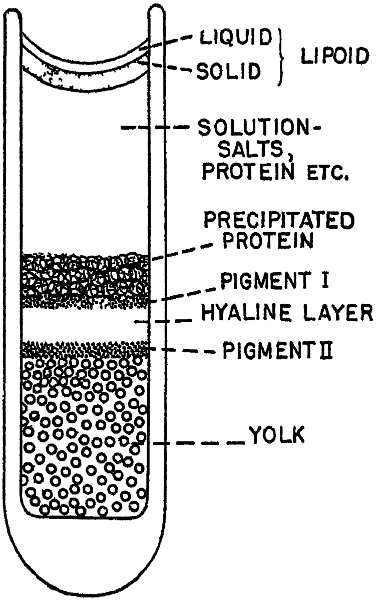


Figure 3. Diagrammatic representation of the stratification of gastrula explants after high-speed centrifugation.

the abscissa at zero. The fact that it does not do so presumably means that the "active" fraction still includes substances, such as lipoidal material and pigment, which are metabolically inert. It is interesting to note that, when respiration is corrected for the amount of inactive material indicated by the graph, the respiratory rates of different parts of the gastrula are approximately identical. It may be mentioned in passing that the inability to obtain, by direct means, a measure of the metabolically active fraction of a piece of tissue emphasizes again the inadequacy of the usual units such as dry weight, total nitrogen, tissue volume, etc., as quantitative indices of the amount of embryonic material used in comparative tests of this kind.

Table 4

Yolk Content of Various Regions of the Gastrula and Respiration of "Active Material" of the Tissue

Region	*No. determinations*	*Per cent yolk*	*Per cent active material*	*Resp.*
Dorsal lip	7	56	44	1.8
Presumptive neural plate	7	33	67	7.3
Anterior ectoderm	7	31	69	6.5
Posterior ectoderm	7	43	57	5.2
Yolk endoderm	8	66	34	3.8
Right lateral ectoderm	2	34	66	4.7
Left lateral ectoderm	2	34	66	5.1
Chorda	2	63	37	3.3

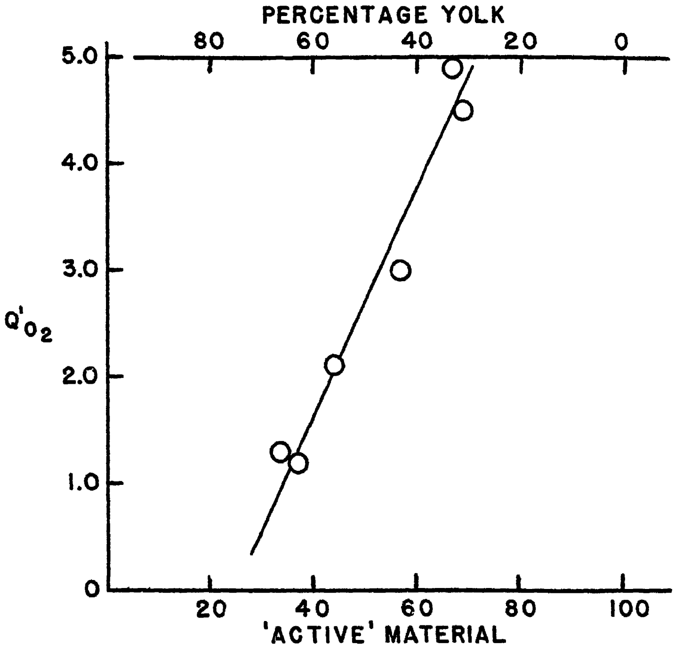


FIGURE 4. Graph showing a plot of the respiratory rate (mμl. O_2/μg. N/hour) of explants of various regions of the gastrula against the percentage yolk content of the regions.

The data summarized in TABLES 2 and 4 are of particular interest when considered in relation to Child's theory of axial gradients as it applies to organizer action. From the results of various non-quantitative tests (such as differential susceptibility to toxic substances, differential staining, etc.) which have been applied to the amphibian gastrula or to other developing or regenerating animals, Child (1929, pp. 50–51) has concluded that ". . . at the time of gastrulation the region about the dorsal lip is the region of most intense metabolic activity in the embryo and that its dominance as an organizer results from this condition." In addition, he has stated ". . . that various facts indicate a very great increase in metabolism during early stages of amphibian development up to the gastrula and that the dorsal lip region at the time of gastrulation is probably not only the region of most intense metabolism in the embryo at that stage, but also of most intense metabolism in embryonic development, because it is physiologically youngest at that stage." In other words, if these statements are interpreted correctly, Child contends that the organizer is an organizer "primarily because of its quantitatively greater activity" as compared with other parts of the embryo. More recently, the same author (1941) has again expressed this view in concluding that "in general, the natural inductors are, or at certain developmental stages, become, high gradient levels, and their inducing action apparently consists in alteration of conditions in regions representing lower levels. The inductor apparently represents a higher range of gradient levels and by raising the level of presumptive neural plate, makes possible further development." Barth (1942) has shared the opinion that "the stimulus for organization proceeds from tissues exhibiting a high rate of oxygen consumption."

Needham (1942), on the other hand, has expressed the opinion that "although gradients of susceptibility to lethal agents, intensity of vital staining, etc., had been clearly demonstrated in a large number of adult invertebrates and many embryos, no evidence whatever has been brought forward to justify a belief in the existence of 'respiratory' or 'metabolic' gradients in embryos, and very little satisfactory evidence for this in the case of adult organisms." Furthermore, after reviewing the available data on the subject, Needham concludes that there is little basis for considering that respiratory gradients are in any way connected with the phenomenon of induction in the amphibian gastrula. Child (1946), in a further attempt to support his view, has stated, however, that the data on respiratory determinations of various gastrular regions "are, at present, far from agreement with Needham," and he rejects the determinations of Boell and Needham (1939), and presumably all others which fail to show a difference in the respiratory rates of dorsal lip tissue and ventral ectoderm, as being without definite significance since no account was taken of varying amounts of yolk in the pieces used in comparative tests. "The dorsal lip cells," he writes, "contain more yolk than those of ventral ectoderm. If they respire at the same rate without correction for yolk, the metabolizing protoplasm of the dorsal lip must actually respire at a considerably higher rate than that of ventral ectoderm." It is clear, however, from an examination of the data in TABLE 4, that even after correction for yolk content, the dorsal lip does not have a higher respiration than ventral ectoderm. A more significant comparison, from the standpoint of Child's theory, would be that between dorsal lip and presumptive neural plate. However, the data in TABLE 4, as well as those of Barth (1942), provide no support for the view that these regions of the gastrula represent respectively "higher" and "lower" gradient levels in the developing system. Furthermore, it would seem unlikely that a quantitative level of metabolism, or difference of level between inducing and induced tissues, is necessary, since it has been experimentally demonstrated that induction and neural differentiation can occur in the presence of cyanide in a concentration sufficient to inhibit respiration by 80 to 90 per cent (Brachet, 1939; Barnes, 1944; Boell, unpublished experiments) and under anaerobiosis (Brachet, 1939). It is of further interest to note that Philips (1942) found no difference of significance in the respiratory rates of various presumptive regions of the early chick blastoderm. Moreover, Lindahl and Holter (1940) reported that the respiratory rates of animal and vegetal halves of sea urchin embryos (*Paracentrotus lividus*) were the same in spite of their different developmental potentialities. Thus, there can be little doubt that inductive action is not associated with quantitatively higher respiratory metabolism in the dorsal lip than in the presumptive neural plate or in other parts of the gastrula.

Respiratory Quotients and Glycolysis. Studies by different workers on the respiratory quotients of various gastrula regions are in complete agreement in showing a value of approximately unity for the dorsal lip tissue and considerably lower values for ventral ectoderm. It has already been noted that Brachet observed, in comparative tests on dorsal lip material and ventral ectoderm, a greater difference in favor of the dorsal lip for carbon dioxide output than for oxygen uptake, and this was correctly interpreted as indicating a difference in the respiratory quotients of the two regions. Subsequent work, summarized in TABLE 5, has abundantly confirmed Brachet's initial observation.

TABLE 5

SUMMARY OF RESPIRATORY QUOTIENTS OF DORSAL LIP AND VENTRAL ECTODERM FROM THE AMPHIBIAN GASTRULA

Investigator	*Material*	*R.Q. DL*	*R.Q. VE*
Brachet	*Discoglossus*	1.02	0.73
Brachet	*R. fusca*	0.97	0.80
Boell, Koch, & Needham	*A. mexicanum*	0.98	0.87
Needham, Rogers, & Shen	*R. temporaria*	0.92	0.81

Brachet (1934) has shown that the R.Q. of intact amphibian eggs averages approximately 0.66 during cleavage stages but rises abruptly to approximately 1.0 during gastrulation, at which point it then remains during subsequent development. Mendes (1948, in press) has extended the study of the respiratory quotient during development of the frog, *Rana pipiens*, to include all stages from immediately after fertilization to the end of the pre-feeding period. He has found that the R.Q. of newly fertilized eggs and of embryos in the early cleavage stages is unity. However, in the late blastula stage, the R.Q. drops to an average value of about 0.7, only to rise again during gastrulation.

Boell, Needham, and Rogers (1939) have shown that the rise observed by Brachet (1934) and Mendes occurs also in isolated regions of the embryo but values of unity are reached by the dorsal lip region in advance of the ventral ectoderm. In the case of the former tissue, the respiratory quotient becomes unity as soon as gastrulation commences, suggesting that the process of invagination is in some way associated with carbohydrate metabolism. Gastrulation in the ventral tissues, presumably, is also concerned with carbohydrate breakdown and an R.Q. of 1.0, but the process appears later in development than it does on the dorsal side of the embryo and it is probable that the respiratory quotient of ventral ectoderm does not reach unity until the tissue becomes underlain by mesoderm.

These respiratory quotients, although admittedly difficult to interpret, have generally been regarded as indicating that carbohydrate metabolism predominates in the dorsal lip region and that ventral ectoderm shows a mixed metabolism. It should be recognized, however, that complete combustion of protein with ammonia as the nitrogenous end product will likewise yield a respiratory quotient near unity. However, it seems unlikely that this is the case in the dorsal lip region of the gastrula since it has been shown that very little ammonia is produced under aerobic circumstances by explants of dorsal lip (Brachet, 1939; Boell, Needham, and Rogers, 1939; Needham, Rogers, and Shen, 1939).

The conclusion which has been derived from the respiratory quotient data reported above, *viz.*, that gastrulation in some way involves carbohydrate breakdown, is made more tenable by direct determination of glycogen in various gastrular regions. Woerdemann (1933) first showed, by a histochemical method, that glycogen disappears from the dorsal lip tissue during invagination. Although the technique used in obtaining this result was not above criticism, Woerdemann's observation was confirmed by Heatley (1935) and Heatley and Lindahl (1937). These authors compared the glycogen content of corresponding regions in the blastula and gastrula of the axolotl embryo and found a significant decrease (31%) in glycogen only in the dorsal lip region (see TABLE 6).

TABLE 6
TOTAL GLYCOGEN IN THE AXOLOTL EMBRYO
Figures in mg. % dry weight
(From HEATLEY & LINDAHL, 1937)

Region	*Blastula*	*Gastrula*	*Per cent decrease*
D 1	17.8	16.5	7
D 2	12.0	8.3	31
D 3+V 3	4.3	3.9	9
V 1	16.7	16.5	1
V 2	10.0	9.3	7

Recently, Jaeger (1945) has confirmed this finding in analyses of *Rana pipiens* embryos (TABLE 7). It is interesting that Jaeger found approximately the same percentage decrease in glycogen (33%) in the dorsal lip

TABLE 7
TOTAL GLYCOGEN IN REGIONS OF FROG EMBRYO
Figures in mg. % dry weight
(From JAEGER, 1945)

Region	*Before gastrulation*	*After gastrulation*	*Per cent decrease*
DL *in situ*	12.3	8.2	33
VL *in situ*	14.0	10.1	28
DL explant	11.2	11.1	1
VL explant	10.8	10.3	4.5
DL+ECT explant	12.8	12.5	2.5

region as was observed by Heatley and Lindahl. Unfortunately, her analyses did not include a comparison with other parts of the gastrula during the same developmental period, but one may conclude, on the strength of Heatley and Lindahl's observations, that the decrease in other regions is perhaps small. Jaeger went on to show, in confirmation of Tanaka's (1934) earlier report, that glycogen was likewise lost from the ventral lip when it underwent invagination. Apparently, glycogen loss from these regions was not an "autonomous" process, for explants of dorsal and ventral lip in saline solutions showed no appreciable decrease in glycogen. Furthermore, glycogen was not lost from presumptive dorsal or ventral lip material in hybrid embryos in which developmental arrest occurred at the gastrula stage. Nor was glycogen loss associated with the induction of neural tissue, since no decrease of glycogen was observed in explants of dorsal lip and competent ectoderm in which neural differentiation had occurred (as indicated by histological appearance of the tissues). Glycogen loss seems not to be associated simply with the assumption of an "internal position" of the tissue after gastrulation, for Raven (1935 a and b) demonstrated that pieces of ectoderm implanted into the blastocoel of the amphibian embryo showed no decrease in their glycogen content. However, a piece of dorsal lip material implanted so as to go through its normal invaginative movements did lose glycogen. Furthermore, a piece of ectoderm implanted in the dorsal lip region and carried into the interior of the embryo by the invagination of the dorsal lip lost glycogen in the same way as did the dorsal lip. From these observations, as well as from those of Jaeger, it appears highly probable that glycogen disappearance is in some way connected with the changes in cell shape and with cell movements during gastrulation. It may be as Needham (1942, p. 191) has suggested, that more than coincidence is involved between the contractile appearance of the blastoporal cells and their utilization of glycogen.

It is noteworthy that Jaeger's figures for dorsal lip tissue were obtained by comparing corresponding regions in embryos at stages 10 and 12, and that those for ventral lip were obtained on stages 11 and 14. Jaeger gave no indication of the time required for development between these stages, but reference to Shumway's (1940) time table for normal development in *Rana pipiens* indicates that the time interval between stages 10 and 12 is 16 hours, whereas 28 hours are required for the embryo to advance from stage 11 to stage 14. Since during these times the quantitative decrease in glycogen was about the same in both regions, the rate of glycogen loss per unit of time is correspondingly greater, indeed almost twice as great, in the dorsal lip than in the ventral lip tissue. This result is interesting in comparison with the data of Boell, Needham, and Rogers (1939) showing that anaerobic glycolysis occurred both in dorsal lip and in ventral tissues of the embryo, but that the rate of glycolysis was about three times greater in the dorsal lip than in ventral ectoderm (TABLE 8).

TABLE 8

ANAEROBIC GLYCOLYSIS ($Q_L^{N_2}$) IN DORSAL LIP AND VENTRAL ECTODERM OF THE AMPHIBIAN GASTRULA

	DL	*VE*
Rana temporaria	+0.63	+0.21
Triton alpestris	+0.42	+0.11

The Development of Enzyme Systems in the Embryo

The organic matter of the amphibian egg may be divided into two fractions. One of these is the protoplasmic or metabolically active portion; the other consists of storage material or yolk, and, as was indicated in the foregoing discussion, there is evidence that this fraction is relatively inert metabolically and serves mainly as a source of raw material for development. It is worthy of note that, during the period of development covering the first 600 hours after fertilization, *i.e.*, during the period of egg and larval development included in Harrison's normal stages, the embryo takes in no food. The total organic matter of the egg thus remains essentially constant; actually it decreases slightly as can be seen in data for dry weight (Dempster, 1930, 1933) or total nitrogen (Wills, 1936). However, during development, the relative proportion of protoplasmic material and yolk is constantly changing as the yolk becomes gradually transformed into new protoplasm and the products of protoplasmic activity. It would seem that the raw materials of the egg represented in the yolk are present in highly concentrated form and that their transformation into protoplasm involves considerable hydration (Gray, 1926; Dempster, 1930, 1933; Boell, 1945). Indeed the embryo of *Amblystoma punctatum* takes in, during its pre-feeding development, approximately 20 mg. of water, and the dry weight is reduced accordingly from 35 per cent at stage 18 to less than 10 per cent at stage 45.

Respiration of the Developing Amphibian Embryo. Unfortunately, we know nothing about the rate at which yolk is being transformed into protoplasm in the normal development of the amphibian embryo, since so far it has been impossible to determine the quantities of the two entities separately. However, an indirect indication of the process may be seen in the changing metabolic activity during development of the embryo as indicated by the respiratory rate (*cf.* Gray, 1927, 1929). A graphic summary of data relating oxygen consumption of *Amblystoma punctatum* to the developmental age of the embryos is seen in FIGURE 5. It is apparent, from the graph, that the amount of oxygen consumed by the embryo increases smoothly and steadily throughout development. Abrupt changes in the rate of respiration, or cyclic variations such as appear in the development of the sea urchin (Lindahl, 1939) or the grasshopper (Boell, 1935) are conspicuously absent. There is no indication of increased energy expenditure for such morphogenetic processes as gastrulation,

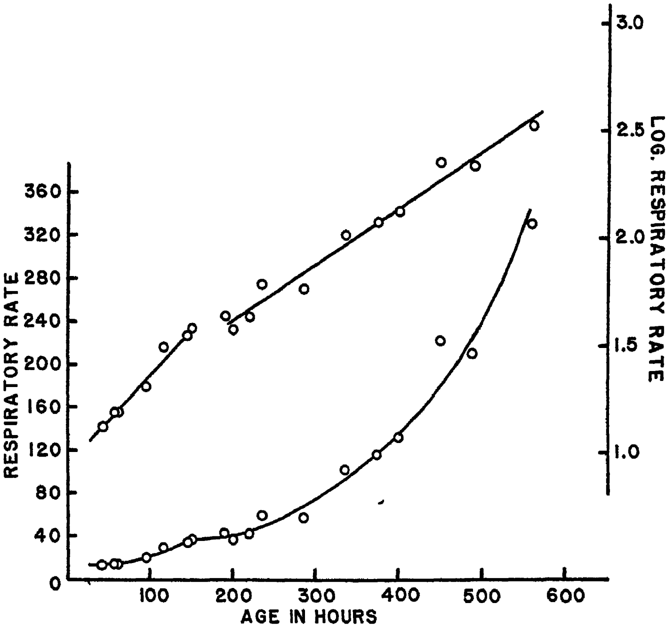


FIGURE 5. Graph showing the respiratory rate of intact embryos of *Amblystoma punctatum* during development. The arithmetic scale on the left relates to the lower curve; the logarithmic scale on the right, to the upper curve. Respiration is expressed as mμl. O_2/100 μg. dry weight/hour.

the differentiation of the primary germ layers or their further elaboration into the definitive tissues and organs of the embryo. Obviously, any energy expenditure for these processes—since they themselves occur gradually—is merely represented as part of the general rise in oxygen consumption which occurs during development. Similar results have been obtained for a number of other amphibian species by Wills (1936), Atlas (1938), Fischer and Hartwig (1938), Hopkins and Handford (1943), Moog (1944), Barnes (1944), Spiegelman and Steinbach (1945), and Barth (1947). It may be concluded, then, that increase in respiration is a reflection of the progressive growth in the embryo of metabolically active material.

Growth is an exponential process (*cf.* Brody, 1945, chapter 16) and obeys the general relation $X = a.e^{kt}$, or $\log X = \log a + kt (\log e)$. From this, it follows that any data which obey the equation should fall on a straight line when the logarithm of the growing entity is plotted against time. If respiratory increase during development is due to the growth of protoplasm, it would be expected that the course of respiration should increase exponentially. Curve *B* of FIGURE 5 indicates that this expectation is met. It is apparent in the curve that the rate of respiratory increase is not constant during development but occurs in two cycles with a "break" appearing on about the sixth or seventh day of development, *i.e.*, at approximately stage 32 or 34. The values of k for the two cycles of growth are, respectively, 0.0098 and 0.0060 (see TABLE 9).

TABLE 9

TABLE OF CONSTANTS FOR THE GROWTH OF VARIOUS ENZYMES AND BIOCHEMICAL ENTITIES IN THE EMBRYOS OF *AMBLYSTOMA PUNCTATUM*

	k
Respiration—40 to 180 hours	0.0098
Respiration—180 to 560 hours	0.0060
Cytochrome oxidase*—60 to 560 hours	0.0075
Cytochrome oxidase†—115 to 630 hours	0.0075
Succinic oxidase—115 to 630 hours	0.0061
Cholinesterase—200 to 400 hours	0.0196
Cholinesterase—400 to 600 hours	0.0092
Volume of central nervous system—200 to 560 hours	0.0006

* *p*-Phenylenediamine used as substrate. Determinations in Warburg apparatus.
† Ascorbic acid used as substrate. Determinations in Cartesian diver apparatus.

Atlas (1938), Moog (1944), Barnes (1944), and Spiegelman and Steinbach (1945) have reported a similar situation in the respiration of *Rana pipiens* and *Rana sylvatica*, but in these species the break occurs after the end of gastrulation, that is, much earlier than in the salamander embryo. A semilogarithmic plot of Barth's (1947) data for *Rana pipiens* similarly shows a break at about stage 15 (Shumway's table). The older data of Bialascewicz and Bledovski (1915, data taken from Needham, 1931, p. 677) likewise can be expressed by the equation mentioned above. In this work, the inflection in rate occurs at about the time when the external gills make their appearance, *i.e.*, at about 100 hours of development.

A semilogarithmic plot of the data for *Amblystoma tigrinum* (Hopkins and Handford, 1943) and *Amblystoma mexicanum* (Fischer and Hartwig, 1938) shows a break at about the same stage as in *Amblystoma punctatum*. Thus, it appears that a difference in the appearance of the break is characteristic of Urodeles and Anurans. It would be of interest to investigate this point further. It may be mentioned, in passing, that a similar inflection in the respiratory rate of the chick embryo can be seen when the data of Murray and Hasselbalch (Needham, 1931, Figure 156, p. 717) and Romanoff (1941) are plotted on a semilogarithmic scale. In the chick, the change in rate occurs at about the seventh day of incubation.

The fundamental cause of the break in the respiratory curves for these forms remains obscure, although it can hardly be doubted that the changing slopes of the curves reflect in some way the rate of transformation of yolk into active protoplasm. It is of considerable interest that in *Amblystoma* the change in rate of respiratory increase coincides with the time when the embryo begins to deviate markedly from the spherical form, and occurs only slightly before the initiation of the heartbeat and the establishment of circulation. Perhaps the rate of respiratory increase is limited to some extent by the development of respiratory surface as well as by the maturation of the circulatory system as an oxygen carrier and as a means of transporting raw materials from the yolk mass to the rest of the embryo.

Cytochrome Oxidase. It has been shown by a number of investigators that the amphibian egg is highly sensitive to cyanide (Brachet, 1934; Barnes, 1944; Boell, 1945), and it has been generally concluded that respiration in the embryos of Amphibia proceeds by way of the cytochrome-cytochrome oxidase system. Measurements of the cytochrome oxidase activity of *Amblystoma punctatum* embryos (Boell, 1945) show that there is sufficient activity of this enzyme during all stages of development to account for the respiration of the embryos. FIGURE 6, curve *A*, illustrates that the increase in cytochrome oxidase activity during development follows a course similar to that of respiration, and like respiration increases exponentially with time. The similarity in the shapes of these curves, as well as the marked sensitivity to cyanide, thus suggest strongly that respiration is largely mediated by the cytochrome-cytochrome oxidase system during development. However, the concentration of cytochrome oxidase apparently is not the factor which limits the rate of increase in respiration, since k for growth of cytochrome oxidase is 0.0075 while for respiration k is 0.0060.

Spiegelman and Steinbach (1945) have reached a similar conclusion from a study of cytochrome oxidase in *Rana pipiens* embryos. These authors report that, although increase in respiratory rate occurs, cytochrome oxidase activity remains uniformly high throughout development and that enzyme synthesis apparently does not take place. This result is somewhat surprising, for it suggests that the unfertilized egg contains the full complement of cytochrome oxidase necessary for the embryo throughout development and that during the process the enzyme merely becomes distributed to the various cells and tissues of the growing organism. Furthermore, synthesis of the enzyme has been demonstrated in the embryos of other forms, as, for example, the grasshopper (Bodine and Boell, 1935; Allen, 1940) and the chick (Albaum and Worley, 1942; Albaum, Novikoff, and Ogur, 1946). Moreover, it may be mentioned that Spiegelman and Steinbach made measurements on embryos only up to stage 19 of Shumway's (1940) table which corresponds approximately with Harrison's stage 34 or 35 for *Amblystoma punctatum*. In *Amblystoma*, the major increase in enzyme activity occurs after this stage, and unpublished experiments of the author indicate that an increase in cytochrome oxidase occurs during the comparable period of development in *Rana pipiens* as well. These experiments show, in addition, that determinations of cytochrome oxidase in early stages of *Rana pipiens* are complicated by the fact that reducing substances in the egg homogenates apparently interfere with the oxidation of *p*-phenylenediamine (compare Brachet, 1945, p. 350), and the full activity of the enzyme may not be obtained by the usual methods. It is interesting to note that Brachet's (1934) paper shows that p-phenylenediamine actually decreased slightly the oxygen consumption of homogenates of *Rana temporaria* eggs.

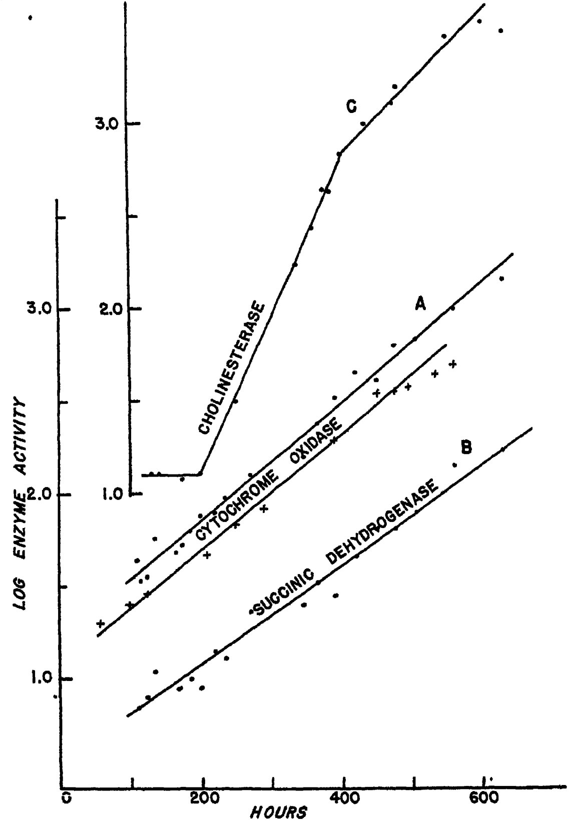


FIGURE 6. Semilogarithmic plot of enzyme activity against age of embryos of *Amblystoma punctatum*. The right-hand ordinate relates to cholinesterase activity, the curve for which has been shifted vertically to avoid superposition of curves. Enzyme activity is expressed as mμl. gas produced or consumed in the Cartesian diver apparatus/100 μg. dry weight/hour.

Succinic Oxidase. FIGURE 6, curve *B*, likewise shows that synthesis of succinic oxidase, as measured by the method of Schneider and Potter (1943), occurs throughout development (Boell, 1946a). Between 100 hours after fertilization and the end of the pre-feeding period (stage 46), succinic oxidase activity increases exponentially with time, the growth constant, k, being 0.0061. This is significantly less than the constant for cytochrome oxidase, but practically identical with that of the respiration of intact embryos during the same period of development. It is of interest that a break, comparable to that for respiration, is not apparent in the growth curves for either of these enzymes.

Schneider and Potter (1943) assume that measurement of the succinic oxidase system by their method actually determines the activity of succinic dehydrogenase. It is fairly certain that the limiting factor in the overall rate of oxygen uptake is neither cytochrome oxidase nor cytochrome, since the former is usually present in excess in the embryo homogenates and the latter can be supplied in excess by the addition of cytochrome *c*, but the method is incapable of distinguishing whether succinic dehydrogenase or some intermediary between succinic dehydrogenase and cytochrome *c* is the limiting factor in the reaction. It will be interesting to have data on succinic dehydrogenase activity alone by the ferricyanide method of Quastel and Wheatley (1938), in order to gain information on this point. When such data are at hand, it may be possible to assess the significance of the identity in growth rates of respiratory and succinic oxidase activities.

Cholinesterase. So far, in this discussion, evidence has been presented in support of the view that increase in respiratory and enzyme activity during development is a reflection, and to some extent a measure, of the production of metabolically active protein from the storage proteins and other materials in yolk. The question naturally arises as to whether the growth of all enzymes may be expected to follow a similar course or whether there is evidence of specialization and individuality in these processes of biochemical differentiation. The results of a study of the development of cholinesterase activity in the embryo of *Amblystoma punctatum* provides interesting information along this line. FIGURE 7, taken from Sawyer's (1943) paper, shows that cholinesterase activity increases progressively throughout the major part of the developmental process. In premotile stages, the increase in cholinesterase activity is slight, but at the time of beginning motility in the embryos it increases abruptly. From a correlation of the development of behavior according to the pattern described by Coghill (1929) and the development of cholinesterase activity, Sawyer was led to conclude that functional maturation of the neuromuscular apparatus of *Amblystoma* coincides with the development of cholinesterase to a quantitatively high level.

The shape of the curve in FIGURE 7 is similar, in some respects, to the curves relating respiration, cytochrome oxidase, and succinic oxidase to time of development. However, closer examination of these curves reveals that they are only superficially similar. We have repeated Sawyer's work in assaying the cholinesterase activity of embryos during development and have fully confirmed his observations. As curve *C* of FIGURE 6 shows, there is essentially no increase in cholinesterase activity during the period of premotile development between 130 and 200 hours (from stage 25 to stage 34). However, at the time when the embryo is first capable of making rapid, repetitive movements, cholinesterase activity is considerable and the enzyme increases in activity throughout the remainder of the period of development studied. The rate of cho-

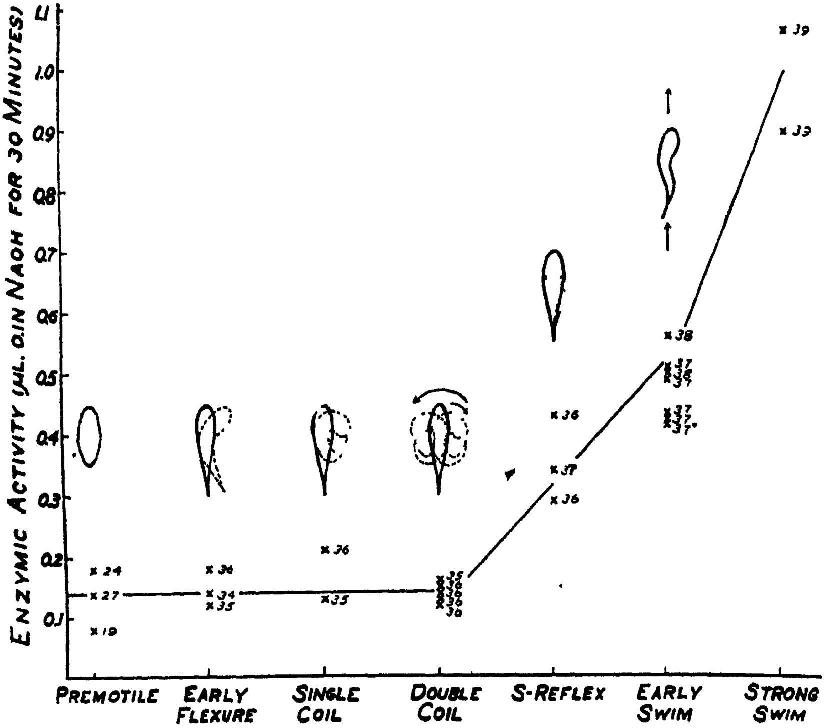


FIGURE 7. Summary of Sawyer's (1943) study of the relationship between the development of behavior and cholinesterase in *Amblystoma punctatum*. The numbers opposite the experimental points denote Harrison's stages.

linesterase synthesis is not constant, however, throughout the entire period. A break in the growth curve occurs after 400 hours or at approximately stage 41–42. Before the inflection, k is 0.0196, afterwards it is 0.0092. It is of considerable interest to note that the time at which the break occurs is correlated with decreased responsiveness of the embryo to mechanical stimulation, as shown in Detwiler's (1946 a and b) studies.

It would appear, from the differences in their growth rates, that the various processes of biochemical differentiation described in the foregoing discussion have certain specific and individual characteristics. If it be true that the increase in activity of respiration and respiratory enzymes reflects an increase in metabolically active material, it seems equally true that the changes in cholinesterase activity do not. The measurements of cholinesterase activity mentioned above were made on homogenates of whole embryos, but there is evidence that the enzyme is associated, at least in early stages, with the development of the nervous system. Sawyer (1943) and Boell and Shen (1944), in comparative studies of the enzyme activity of neural and non-neural tissues, were able to demonstrate a marked localization of the enzyme in the nervous system as early as stage 19 (closing neural folds), and that the difference between neural and non-neural tissues became progressively greater during subsequent development. Furthermore, assays of cholinesterase activity in secondary nervous systems, produced experimentally by the inductive action of dorsal lip tissue (chorda mesoderm) which was implanted into the blastocoele of host embryos, gave unmistakable evidence of an increased concentration of cholinesterase in the differentiating nervous system (TABLE 10; Boell and Shen, 1944). In this study, the results indicated that the cholinesterase values for the secondarily induced nervous system were quantitatively of the same order as those of the primary nervous system of the embryo and considerably higher than those of ectoderm or skin. Thus, these experiments indicate that the phenomenon of induction alters the biochemical as well as the morphological fate of the induced tissue.

TABLE 10

CHOLINESTERASE IN EXTRACTS OF PRIMARY AND SECONDARY NEURAL TISSUES
(From BOELL & SHEN, 1944)

	Q'_{ACH}		
Stage	*Primary nervous system*	*Secondary nervous system*	*Host ectoderm*
26	+32	+38	+22
34	+45	+17	+ 6
34	+66	+58	+14

Q'_{ACH}=mμl. CO_2 produced in hydrolysis of acetylcholine by extract containing 1 μg. N per hour.

During development, the nervous system increases in size, and one is therefore led to inquire whether the increase in cholinesterase activity in the embryo may be simply due to the growth of the nervous system in volume. It is obviously impossible to obtain a measure of all the neural tissues in the embryo, because of the development of nerves and other peripheral elements, but it was thought that some information on this point could be gained by making a study of the growth in volume of the central nervous system. The data were secured by projecting 10 μ serial sections of *Amblystoma* embryos at a magnification of 150 times. The outlines of the central nervous system were drawn on paper and the areas of the tracings determined by means of a planimeter. In the head and trunk region, each section was drawn and analyzed, but in the tail, where the diameter of the spinal cord was fairly uniform over relatively long distances, only every fourth or eighth section was measured. The sum of the planimetric readings was then taken as a measure of the volume of the central nervous system at each stage. Although the shrinking of nervous tissue in fixative may not be uniform at all stages of development, it was felt, nevertheless, that this method of determining the volume of the nervous system in the embryo yields data which are infinitely more accurate than any which could be obtained from dissected and weighed organs. FIGURE 8 indicates that the growth of the nervous system is essentially uniform between stages 34 and 46 and occurs at a relatively low rate; indeed, the rate of growth is only one-thirtieth that of cholinesterase. In addition, inspection of the curves shows that they are qualitatively quite different. It thus appears that the increase in the concentration of cholinesterase is not simply a function of the growth in volume of the central nervous system. No doubt, the increase in enzyme activity parallels, to some extent, morphological and physiological differentiation of the nervous system. However, part of the increase of enzyme content of the whole embryo is unquestionably due to the development of cholinesterase in non-neural tissues of the embryo, for example, in muscle (Sawyer, 1943).

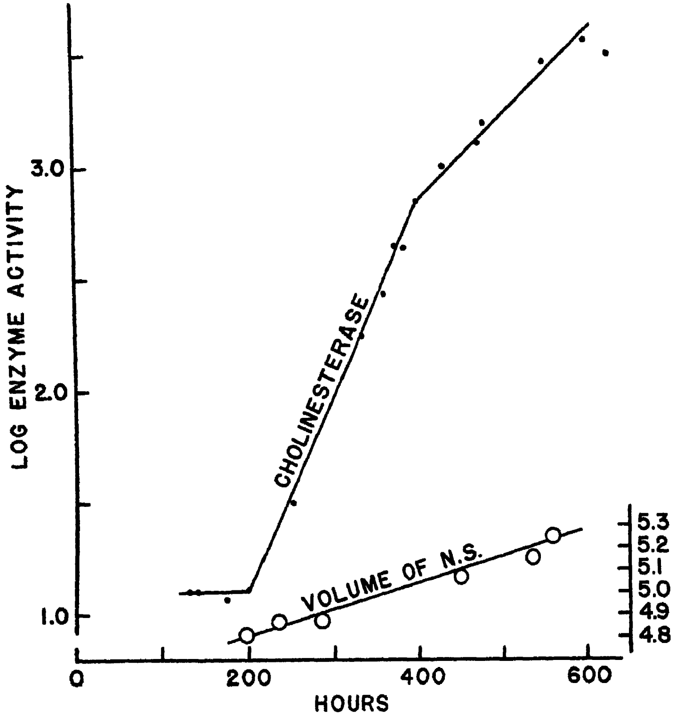


FIGURE 8. Curves showing the relative growth rates of cholinesterase activity and the volume of the central nervous system in *Amblystoma punctatum*. The upper curve is the same as curve *C* in FIGURE 6.

The Effect of Di-isopropyl Fluorophosphate. Di-isopropyl fluorophosphate (DFP) is one of a group of alkyl fluorophosphates studied by the chemical warfare services of both this country and England during the last war. According to Dixon and Needham (1946), DFP was found to inhibit cholinesterase and other esterases but had no effect on the activity of a number of other enzymes. The compound is characterized by its intense anticholinesterasic activity, a concentration as low as 10^{-11}M producing a definite effect. Dixon and Needham conclude that the compounds to which DFP belongs "are the most powerful and specific enzyme inhibitors known." Mazur and Bodansky (1945) reported that the inhibition produced by DFP was irreversible as judged by the inability of serum esterase to regain its activity after prolonged dialysis (however, see Bulloch, Grundfest, Nachmansohn, Rothenberg, and Sterling, 1946). They also showed that regeneration of cholinesterase activity in the nervous systems of experimental animals treated with DFP required an extremely long time. In rabbit brain, for example, cholinesterase activity was only 90 per cent of normal 50 days after treatment.

In view of these properties of the compound, it was felt that treatment of *Amblystoma* embryos with DFP so as to produce what has been aptly termed a "biochemical lesion" might yield interesting information on the functional importance of cholinesterase (Boell, 1946b). It may be recalled that Sawyer (1943) had shown that the behavior responses of *Amblystoma* embryos were depressed in the presence of the reversible inhibitor, physostigmine, and that there was a rough correlation between depression of the enzyme and the loss of normal behavioral ability.

Embryos placed in DFP in premotile stages could apparently be reared indefinitely in a concentration of 0.0001 M with no effect on morphogenesis other than slight retardation of development. However, stronger concentrations were invariably toxic and resulted in the death of the embryos in a few minutes to a few hours. In a few embryos in 0.0001 M DFP, certain abnormalities of development were noted, but in the main the DFP-treated embryos seemed to be completely normal on the basis of external appearance. Moreover, the embryos, even when reared continuously in DFP solutions for a week or more, manifested the usual behavior reactions, and, at the appropriate stages (37–38), responded to tactile stimulation with swimming movements. However, the DFP-treated embryos, although apparently normal in appearance, fatigued more readily than control animals when stimulated, and when tested by Detwiler's (1946 a and b) method for measuring behavior responses quantitatively they gave evidence of greatly reduced activity. After such embryos had been subjected to stimulation, they were usually incapable of responding again until a rest period of considerable length had elapsed. It is interesting to compare these results with the observations of Modell and Krop (1946). These investigators showed that cats that had been treated with DFP were perfectly normal in appearance and were well groomed but on stimulation gave evidence of marked myoneural disfunction. After a few steps, the animals fell over and could not be made to move again even when given a stimulus strong enough to elicit a cry.

Embryos reared in DFP solutions and capable of responding to stimulation were found to possess cholinesterase activity of considerable magnitude, although it was much lower than in control embryos of the same developmental stage. It is of some significance to note that the concentration of cholinesterase in embryos reared in DFP was never less and usually greater than in normal animals at the time when they were first capable of responding to stimulation with rapid, repetitive movements. This is shown in TABLE 11 and would seem to indicate that a considerable excess of cholinesterase over that needed for minimal activity is present in the embryos at most stages (*cf.* Bulloch *et al.*, 1946). Under conditions of stress, the reduced cholinesterase activity may, however, lead to functional abnormality. This is clearly shown in a study of the effects of stimulation on the heart rate of the embryos. When normal embryos are stimulated until they are incapable of responding, the rate of heartbeat usually increases, but continued stimulation of DFP-treated embryos leads to a marked reduction in the rate of heartbeat. However, the beat is again normal after several minutes. Perhaps the residual cholinesterase in the tissues of the treated embryos is insufficient to remove the acetylcholine produced under such conditions of extreme activity. As a consequence, it may diffuse into the blood stream and be carried to the heart where it produces its characteristic depressing effect.

TABLE 11

CHOLINESTERASE ACTIVITY OF *Amblystoma* EMBRYOS
REARED IN 0.0001 M DFP

Days in DFP	*Stage of embryo*		*Cholinesterase activity*		*Response to stimulation*
	When DFP was added	*When tested*	*Control*	*Treated*	
1	25	26	13	0	No motility
2	25	26+	13	0	No motility
—	—	37	32	—	Rapid flutter
—	—	38	63	—	Swimming response
10	22	40	436	65	Embryo swims; tires
7	27	40	436	87	Embryo swims; tires
12	20	41	973	34	Rapid flutter
14	20	42	1100	48	Behavior not noted
11	27	42+	1540	163	Embryo swims; tires
17	22	45−	2880	855	Embryo swims; tires

Apparently, synthesis of cholinesterase can take place in the embryo in the presence of DFP. TABLE 11 shows that the cholinesterase activity of embryos treated with DFP during premotile stages and reared continuously in DFP solutions increases five-to six-fold during the course of a week to ten days. After 17 days in DFP solutions, the cholinesterase activity may be as much as 65 times greater than in premotile embryos. Although the absolute level of enzyme activity is much lower in treated than in untreated animals of the same morphological stage, being on the average only 10 per cent of the control values, the rate of increase with development is approximately the same in the two groups of embryos. Apparently, therefore, development of the enzyme is not interfered with by DFP, although the newly synthesized enzyme is inhibited as soon as it is formed. The fact that the percentage enzyme activity of the treated animals in relation to the controls is approximately the same at all stages of development, suggests that an equilibrium exists between inhibited and uninhibited enzyme.

Summary

The foregoing discussion has dealt primarily with the demonstration of regional biochemical differences in the amphibian gastrula and with the increase in activity of certain enzyme systems throughout embryonic life. Processes of the latter type may be regarded as examples of biochemical differentiation since they represent special aspects of the synthetic or formative activities of the embryo. The discipline of chemical embryology is relatively young, and, although much progress has been made in the field in the past two or three decades, our knowledge of biochemical aspects of embryonic development is still relatively meager. It may be hoped that future work will supply us not only with more information on biochemical differences and differentiation in the embryo, but, in addition, with increased understanding of the way in which biochemical and morphogenetic processes are integrated in the complex series of events which characterize the transformation of the egg into the individual.

Literature Cited

1. Albaum, H. G., A. B. Novikoff, & M. Ogur. 1946. The development of the cytochrome oxidase and succinoxidase systems in the chick embryo. J. Biol. Chem. 165: 125–130.
2. Albaum, H. G., & L. G. Worley. 1942. The development of cytochrome oxidase in the chick embryo. J. Biol. Chem. 144: 697–700.
3. Allen, T. H. 1940. Enzymes in Ontogenesis. XI. Cytochrome oxidase in relation to respiratory activity and growth of the grasshopper egg. J. Cell. & Comp. Physiol. 16: 149–163.
4. Atlas, M. 1938. The rate of oxygen consumption of frogs' eggs during embryonic development and growth. Physiol. Zool. 11: 278–291.
5. Barnes, M. R. 1944. The metabolism of the developing *Rana pipiens* as revealed by specific inhibitors. J. Exp. Zool. 95: 399–417.
6. Barth, L. G. 1939. Oxygen consumption of the parts of the amphibian gastrula. Proc. Soc. Exp. Biol. & Med. 42: 744–746.
7. Barth, L. G. 1942. Regional differences in oxygen consumption of the amphibian gastrula. Physiol. Zool. 15: 30–46.
8. Barth, L. G. 1947. Studies on the metabolism of development. J. Exp. Zool. 103: 463–486.
9. Bodine, J. H., & E. J. Boell. 1936. Enzymes in Ontogenesis. II. The indophenol oxidase. J. Cell. & Comp. Physiol. 8: 213–230.
10. Boell, E. J. 1935. Respiratory quotients during embryonic development (Orthoptera). J. Cell. & Comp. Physiol. 6: 369–385.
11. Boell, E. J. 1942. Biochemical and physiological analysis of organizer action. Growth 7 (Suppl.): 37–53.
12. Boell, E. J. 1945. Functional differentiation in embryonic development. II. Respiration and cytochrome oxidase activity in *Amblystoma punctatum*. J. Exp. Zool. 100: 331–352.

13. Boell, E. J. 1946a. Succinic dehydrogenase activity during the development of *Amblystoma punctatum*. Anat. Rec. (Suppl.) 96: 91.

14. Boell, E. J. 1946b. The effect of di-isopropyl fluorophosphate on the development of behavior and cholinesterase in *Amblystoma punctatum*. Anat. Rec. (Suppl.) 96: 4–5.

15. Boell, E. J., H. Koch, & J. Needham. 1939. IV. Respiratory quotient of the regions of the amphibian gastrula. Proc. Roy. Soc. B 127: 374–387.

16. Boell, E. J., & J. Needham. 1939. III. Respiratory rate of the regions of the amphibian gastrula. Proc. Roy. Soc. B 127: 363–373.

17. Boell, E. J., J. Needham, & V. Rogers. 1939. Morphogenesis and Metabolism: Studies with the Cartesian diver ultramicromanometer. I. Anaerobic glycolysis of the regions of the amphibian gastrula. Proc. Roy. Soc. B 127: 322–356.

18. Boell, E. J., & J. S. Nicholas. 1940. Respiratory rate and yolk content of the amphibian gastrula. Anat. Rec. (Suppl.) 78: 76.

19. Boell, E. J., & S. C. Shen. 1944. Functional differentiation in embryonic development. I. Cholinesterase activity of induced neural structures in *Amblystoma punctatum*. J. Exp. Zool. 97: 21–41.

20. Brachet, J. 1934. Étude du métabolisme de l'oeuf de grenouille (*Rana fusca*) au cours du développement. I. La respiration et la glycolyse, de la segmentation à l'éclosion. Arch. Biol. 45: 611–727.

21. Brachet, J. 1935. Étude du métabolisme de l'oeuf de grenouille (*Rana fusca*) au cours du développement. 3. Métabolisme respiratoire et "centre organisateur" de la gastrula. Arch. Biol. 46: 25-45.

22. Brachet, J. 1936. Le métabolisme respiratoire du centre organisateur de l'oeuf de grenouille (*Rana fusca*). C. R. Soc. Biol. 122: 108.

23. Brachet, J. 1939. V. Le métabolisme protéique et hydrocarboné de l'oeuf en relation avec le problème de l'organisateur. Arch. Biol. 50: 233–267.

24. Brachet, J. 1945. *Embryologie chimique*, 2nd ed. Masson et Cie. Paris.

25. Brachet, J., & H. Shapiro. 1937. The relative oxygen consumption of dorsal and ventral regions of intact amphibian gastrulae, including observations on unfertilized eggs. J. Cell. & Comp. Physiol. 10: 133–144.

26. Bragg, A. N. 1939. Observations upon amphibian deutoplasm and its relation to embryonic and early larval development. Biol. Bull. 77: 268–283.

27. Brody, S. 1945. *Bioenergetics and Growth*. Reinhold. New York.

28. Bullock, T. H., H. Grundfest, D. Nachmansohn, M. A. Rothenberg, & K. Sterling. 1946. Effect of di-isopropyl fluorophosphate (DFP) on action potential and choline esterase of nerve. J. Neurophysiol. 9: 253–260.

29. Child, C. M. 1929. Physiological dominance and physiological isolation in development and reconstitution. Roux's Arch. Entwickl. 117: 21–66.

30. Child, C. M. 1941. *Patterns and Problems of Development*. Chicago Univ. Press.

31. Child, C. M. 1946. Organizers in development and the organizer concept. Physiol. Zool. 19: 89–148.

32. Coghill, G. E. 1929. *Anatomy and the Problem of Behaviour*. The University Press. Cambridge.

33. Daniel, J. F., & E. A. Yarwood. 1939. The early embryology of *Triturus torosus*. Univ. Calif. Publ. Zool. 43: 321-356.

34. Dempster, W. T. 1930. The growth of larvae of *Amblystoma maculatum* under natural conditions. Biol. Bull. 58: 182-192.

35. Dempster, W. T. 1933. Growth in *Amblystoma punctatum* during the embryonic and early larval period. J. Exp. Zool. 64: 495–511.

36. Detwiler, S. R. 1946a. Experiments upon the midbrain of *Amblystoma embryos*. Am. J. Anat. 78: 115-138.

37. Detwiler, S. R. 1946b. A quantitative study of locomotion in larval *Amblystoma* following either midbrain or forebrain excision. J. Exp. Zool. 102: 321–332.

38. Dixon, M., & D. M. Needham. 1946. Biochemical research on chemical warfare agents. Nature 158: 432–438.

39. Fischer, F. G., & H. Hartwig. 1938. Vergleichende Messungen der Atmung des Amphibien-Keimes und seiner Teile während der Entwicklung. Biol. Zentralbl. 58: 567–589.

40. Gray, J. 1926. The growth of fish. I. The relation between embryo and yolk in *Salmo fario.* J. Exp. Biol. 4: 215–225.

41. Gray, J. 1927. The Mechanism of Cell Division. III. The relationship between cell-division and growth in segmenting eggs. J. Exp. Biol. 4: 313–321.

42. Gray, J. 1929. The kinetics of growth. J. Exp. Biol. 6: 248–274.

43. Heatley, N. G. 1935. Distribution of glycogen in regions of the Amphibian gastrula; with a method for the microdetermination of glycogen. Biochem. J. 29: 2568–2572.

44. Heatley, N. G., & P. E. Lindahl. 1937. Studies on the Nature of the Amphibian Organization Centre. 5. Distribution and nature of glycogen in Amphibian embryo. Proc. Roy. Soc. London B 122: 395–402.

45. Holtfreter, J. 1944. Neural differentiation of ectoderm through exposure to saline solution. J. Exp. Zool. 95: 307–343.

46. Holtfreter, J. 1945. Neuralization and epidermization of gastrula ectoderm. J. Exp. Zool. 98: 161–209.

47. Hopkins, H. S., & S. W. Handford. 1943. Respiratory metabolism during development in two species of *Amblystoma.* J. Exp. Zool. 93: 403–414.

48. Jaeger, L. 1945. Glycogen utilization by the amphibian gastrula in relation to invagination and induction. J. Cell. & Comp. Physiol. 25: 97–120.

49. Lindahl, P. E. 1939. Zur Kenntnis der Entwicklungsphysiologie des Seeigeleies. Z. vergl. Physiol. 27: 233–250.

50. Lindahl, P. E., & H. Holter. 1940. Der Atmung animaler und vegetativer Keimhälften von *Paracentrotus lividus.* C. R. Trav. Lab. Carlsberg (Série chim.) 23: 257–288.

51. Mazur, A., & O. Bodansky. 1946. The mechanism of *in vitro* and *in vivo* inhibition of cholinesterase activity by diisopropylfluorophosphate. J. Biol. Chem. 163: 261–276.

52. Mendes, E. G. 1948. Respiratory quotient during the development of *Rana pipiens* (in press).

53. Modell, W., & S. Krop. 1946. Antidotes to poisoning by DFP in cats. J. Pharmacol. 88: 34–38.

54. Moog, F. 1944. The chloretone sensitivity of frogs' eggs in relation to respiration and development. J. Cell. & Comp. Physiol. 23: 131–155.

55. Needham, J. 1931. *Chemical Embryology.* The University Press. Cambridge.

56. Needham, J. 1942. *Biochemistry and Morphogenesis.* The University Press. Cambridge.

57. Needham, J., & E. J. Boell. 1938. Metabolic properties of the regions of the amphibian gastrula. Proc. Soc. Exp. Biol. & Med. 39: 287–290.

58. Needham, J., V. Rogers, & S. C. Shen. 1939. Morphogenesis and metabolism: Studies with the Cartesian diver ultramicromanometer. V. Aerobic glycolysis measurements on the regions of the amphibian gastrula. Proc. Roy. Soc. B 127: 576–583.

59. Parnas, J. K., & Z. Krasinka. 1921. Ueber den Stoffwechsel der Amphibienlarven. Biochem. Z. 116: 108–137.

60. Philips, F. S. 1942. Comparison of the respiratory rates of different regions of the chick blastoderm during early stages of development. J. Exp. Zool. 90: 83–100.

61. Pickford, G. E. 1943. The distribution of dipeptidase in the salamander gastrula. J. Exp. Zool. 92: 143–170.

62. Quastel, J. H., & A. H. M. Wheatley. 1938. Anaerobic oxidations. On ferricyanide as a reagent for the manometric investigation of dehydrogenase systems. Biochem. J. 32: 936–943.

63. Raven, C. P. 1935a. Experimentelle Untersuchungen über den Glykogen-Stoffwechsel des Organisationszentrums in der Amphibiengastrula. Proc. Koninkl. Acad. Wetenschap. Amsterdam 38: 1107–1109.

64. Raven, C. P. 1935b. Über assimilatorische Induktion in der dorsalen Urmundlippe der Amphibiengastrula. Proc. Koninkl. Acad. Wetenschap. Amsterdam 38: 1109–1115.

65. Romanoff, A. L. 1941. The study of the respiratory behavior of individual chicken embryos. J. Cell. & Comp. Physiol. 18: 199–214.

66. Sawyer, C. H. 1943. Cholinesterase and the behavior problem in *Amblystoma*. I. The relationship between the development of the enzyme and early motility. II. The effects of inhibiting cholinesterase. J. Exp. Zool. 92: 1–29.

67. Schneider, W. C., & V. R. Potter. 1943. The assay of animal tissues for respiratory enzymes. II. Succinic dehydrogenase and cytochrome oxidase. J. Biol. Chem. 149: 217-227.

68. Shumway, W. 1940. Stages in the normal development of *Rana pipiens*. Anat. Rec. 78: 139–147.

69. Spiegelman, S., & H. B. Steinbach. 1945. Substrate-enzyme orientation during embryonic development. Biol. Bull. 88: 254–268.

70. Tanaka, S. 1934. Glycogen distribution in amphibian embryos. Proc. Imp. Acad. Tokyo 10: 689-691.

71. Waddington, C. H., J. Needham, & J. Brachet. 1936. Studies on the nature of the amphibian organization centre. III. The activation of the evocator. Proc. Roy. Soc. London B 120: 173-198.

72. Wills, I. A. 1936. The respiratory rate of developing amphibia with special reference to sex differentiation. J. Exp. Zool. 73: 481–510.

73. Woerdeman, M. W. 1933. Über den Glykogenstoffwechsel des Organisationszentrums in der Amphibiengastrula. Proc. Amsterdam Acad. Sci. 36: 189–193.

FORM CHANGES DURING PRE-GASTRULAR DEVELOPMENT

By J. S. NICHOLAS
Osborn Zoological Laboratory, Yale University, New Haven, Connecticut

THE importance of movements of pre-gastrular materials has long been the subject of analysis by embryologists. The earliest observations on eggs have been those of studying cortical movements, movements of the materials which evidence themselves upon the surface of the egg. These movements and rearrangements are mirrored in the distribution of potencies and potentialities of the whole system in the later embryo. As is well known, Roux's (1885) early studies indicated a concept of fixity within the egg recently employed by Hamburger (1947) in an attempt to give a definite description of prelocalization at the time of fertilization or very shortly thereafter. It has been known since the time of Roux that, in certain forms, fertilization by means of an artificial insemination caused the formation of the grey crescent at a point opposite the region of sperm entrance; that the original cleavage bisects the grey crescent and coincides with the longitudinal axis of the embryo. This fixity has since been challenged and has been found not to exist in many of the other forms. It can be said with safety that there is no absolute fixed relation between the entrance of the sperm, the first cleavage, and the future axis of the embryo. While these coincidences may occur occasionally, according to statistical probabilities, they are not an interlocking determinant series.

The cortical movements of the egg have been most recently studied by Holtfreter (1944). He finds that, even in the unfertilized egg, one may have many of the cortical movements which are so characteristic of the fertilized egg. The movements of the pigment, the characteristic stranding of the pigment particles, the dispersion of pigment in certain areas —all these are obtained in unfertilized eggs. There may even be an attempt at a pseudo-gastrulation in which the surface parts of the egg, without having undergone any division at all, form lips similar to those which are found in gastrulation of a normal, fertilized type. It is true that these will not carry further, and since the material is still in a much more homogeneous condition, the materials of development are not specifically localized for further development of the organism. Holtfreter has repeated and amplified the explanations of Rhumbler (1902) in attempting to correlate the early rearrangements with the later surface movements which are so important in the allocation of materials.

Recently, there has been a renewal of the study in analyzing the movements and positions of parts in the interior of the egg. Daniel and Yarwood (1939) studied the relationship of parts in the various stages of the egg beginning with the ovarian egg, then the coelomic egg, the oviducal egg, and finally the zygote. In each of these, there is a definite stratification of the yolk. The animal pole contains smaller granules than are found in each of the succeeding fifths, until one approaches the vegetal pole where there is an increasing aggregation of the smaller particles. This is in distinct violation of our generally expressed ideas that the vegetal pole of the egg contains all the large yolk-laden elements practically devoid of pigment, and is filled with yolk platelets from which they gain the distensibility of the membrane around the yolk packets. One fails to consider that the endodermal cells at the vegetal pole of the egg are vital, tangible, and mobile elements in the picture of later development. If one surveys the figure which is given by Daniel and Yarwood (1939), it is evident that the median three-fifths of the egg are the ones which contain the heavier amounts of the yolk particles, that the upper fifth and the lower fifth, that is, the animal pole and the vegetal pole, are much less heavily laden with the yolk material. It is particularly so in the zygote stage when fertilization has taken place and the materials are beginning to separate in a much more definitive fashion than they have in the ovarian, coelomic, or oviducal egg.

Schechtman (1934, 1935, 1937), in studying ingression, was the first to notice that material could pass from the outside of the egg in the region of the vegetal pole toward the inside of the egg and, if stained in the uncleaved egg, the stains placed upon the outside to mark the regions which ingress appeared in a columnar form extending to the floor of the blastocoele. This observation is of great importance, for it shows the movement of materials from the outside to the inside of the egg. It also shows their definite alignment along a certain, particular portion of the egg. It shows, in addition, that the movement is limited to rather early stages, since in the early blastula the stain, placed as a spot on the vegetal pole of the uncleaved egg, extends from the outside to the floor of the blastocoele, but if the stain is similarly placed in the late blastula, it appears only in the lower third of the blastular floor. This observation of Schechtman is so striking that it seemed necessary to repeat the staining experiments using different degrees of stain and following the material through to a later stage of development. The phenomena of ingression are proved definitely by Schechtman's work. Its import, however, was not clearly realized by him, and since the interpretation of the significance of the cells found in the ingressed mass must remain for future study, it is necessary to repeat and amplify Schechtman's work in order to find out exactly where the cellular elements bearing the stain are forming. For this reason, the stains were applied to the outside of the egg in a manner similar to Schechtman's, beginning with the unfertilized egg and running a series similar to his. The results are quite interesting in that they show that, if an intense stain is placed upon the uncleaved egg, one can perceive as development continues that there are two lines by which the stain is sent to the inside of the embryo. The first is the one which Schechtman found and which characteristically occupies the central plane of the egg. It is as though one had a central axis of polarity in the amphibian egg due to the ingression of materials which were originally on the outside. It forms a distinct entity extending upward toward the blastocoele as a definite cone of stained material.

At the time of the late blastula, there are certain characteristic changes in organization of the interior which bring about a change in the axes of the embryo. At this time, one finds that the secondary part of the stain which had remained on the outside of the egg is carried around to the ventral lip of the blastopore, and then progresses with the ventral lip of the blastopore across the floor of the gastrocoele. The primary ingression cone has left its imprint on the center of the floor of the blastocoele and, as the gastrocoele encroaches upon the blastocoele, we find that the two colored areas approach each other as the blastocoele begins to lose its form and undergoes definitive progression toward amalgamation with the gastrocoele. In *Amblystoma punctatum*, the blastocoele remains and does not become continuous with the gastrocoele, but the stain which has come into the embryo from the ventral lip of the blastopore progresses forward in the floor of the gastrocoele until it unites with the stain which has come in through the primary ingressive mechanism.

If one applies stain to the embryo during the late blastula stage, the effect of the primary ingression with the cone of material extending inward toward the blastocoele floor is completely lost and one secures only the secondary migration of material which comes in through the ventral lip of the blastopore to the gastrocoele floor. It has been found (Nicholas, 1945) that the material which comes in through the primary ingression lies in the anterior region just under the floor of the foregut marking the region between the heart and the liver in the early stages. This is an important location and probably marks an area in which induction effects are taking place. In recapitulation, then, the staining experiments show definitely that Schechtman's primary ingression is a definite entity to which can be added, with an intensification of the stain, a secondary situation by which the stain will progress through the ventral lip of the blastopore and later join with those fragments of stain which are left from the primary ingression. Recently, Nieuwkoop (1946) has repeated Schechtman's experiments and shown the secondary relationship in an exceedingly clear way. Nieuwkoop interprets the secondary portion of the ingression which comes in through the ventral lip of the blastopore as lying in the surface of the midgut. The primary ingression coincides with the material which Schechtman found in *Triturus* and Nicholas in *Amblystoma*.

It is clear that polar orientation is influenced by the early cortical movements which, in themselves, are a reflection of the reorientation of the materials within the egg. Polar orientation has been said to occur in the ovary (Child, 1941), and undoubtedly in the amphibian egg this

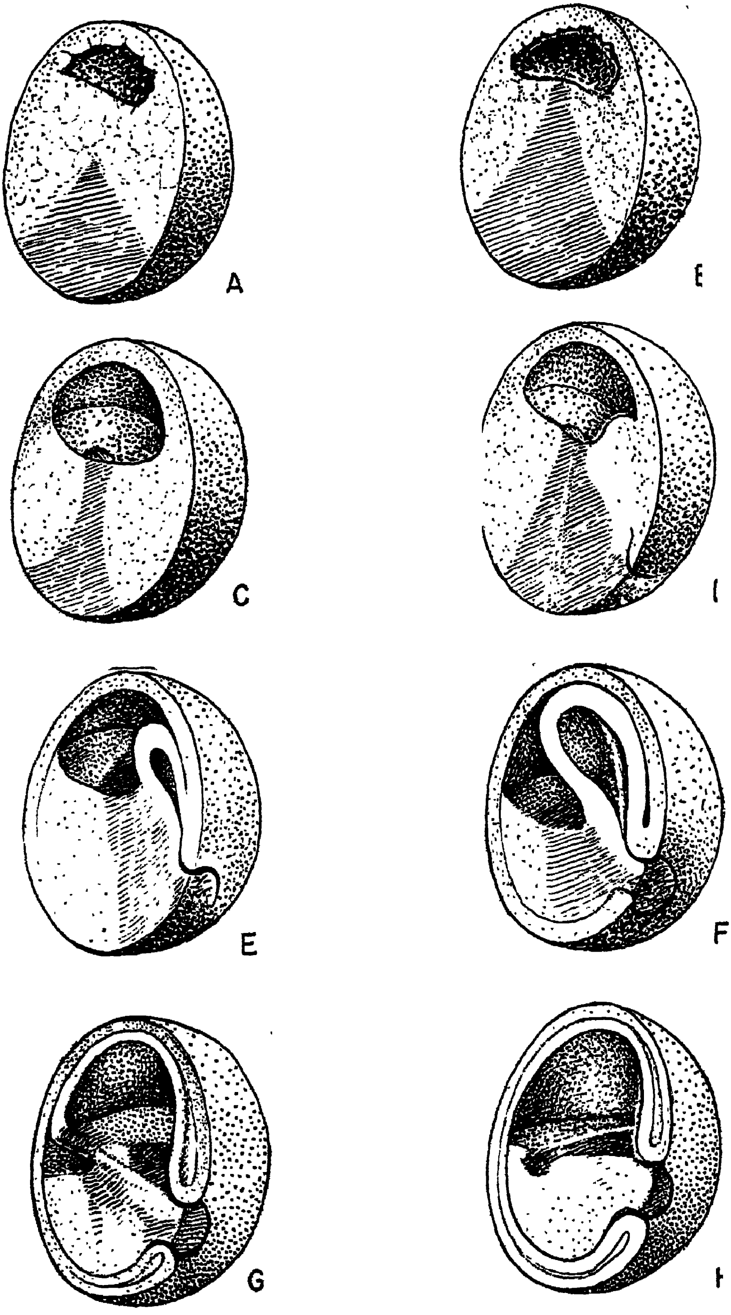


FIGURE 1. (*For description see facing page*)

is partially correct. There is, however, a marked deviation in the animal-vegetal axis shortly before fertilization and continuing afterwards. This may result in the marking of the axis 30 degrees from the original line at the time that the egg was laid. This reaction has been neglected to a large extent because we have not observed the difference in the polarity of the egg before fertilization and afterwards, but if one watches the egg one can see the change in the polarity of the animal-vegetal axis shortly after fertilization and to some extent before then.

The ingressive phenomena are essentially those which lead to the orientation of the mass within the egg after the early fertilization change in the polar axis. The inertness of the egg material which had formerly been thought to be burdened down by yolk has been found to be a fallacy in that an active movement of particles and materials does occur throughout the early stages before segmentation and apparently determines, to some extent at least, the relationship of the cortico-vegetal orientation.

The converse of this type of reaction is seen in the capacity of the embryo to undergo a tremendous amount of development without any of its yolk taking part in the reaction. This is true if one takes into account the fact that the yolk is undergoing a development of its own and has been in contact with the ectodermal and mesodermal portions of the system during their development. Nieuwkoop (1946) removed all of the yolk in the early neurula stage of amphibian embryos. When all of the yolk mass is removed from the neurula, it leaves the mesoderm which has moved inside through the blastopore, the entire chorda mesoderm which has underlain the nervous system, and the lateral mesoderm in its normal position. All of these migrations have been completed before the stage of the neurula. When the yolk is removed in this way, an embryo develops which is deficient in the head region, lacks gills, stomodaeum, all of the intestinal tract and its derivatives, and also the heart. Most of the nervous system, the lateral body wall and musculature are intact, but there is no splanchnopleure whatsoever. That is, the material which was destined to form smooth musculature of the gut is not formed. This speaks for a strong dependence of smooth musculature formation upon the gut structures which it normally surrounds. Another interesting phenomenon occurring after this operation is that two full sets

FIGURE 1. A (stage 5). The stain placed upon the vegetal region of the uncleaved egg extends upward in a cone, which in section appears as a wedge, with its apex projecting into the upper half of the blastula. In B (stage 7), the stain has reached the floor of the blastocoele, and in C is beginning to spread over the floor surface, first as a small circle of stained material, later, stage 8, in a radiating fashion or in a complete coverage of the floor, depending upon the intensity of original staining. During the latter part of stage 8 (D), there is a tendency to have a separation of the stain into a primary segment which passes toward the blastocoele cavity and a secondary aggregation which is drawn toward the presumptive yolk plug. This reaction is more marked in E (stage 9) and F (stage 10) when the folding over of the blastocoele floor and the formation of the underside of the gastrocoele floor begin to have a common cell boundary. The two movements, the primary and secondary, are still separate in stage 11, the stain of the primary being concentrated in the walls of the blastocoele and extending through the cellular boundary into the floor of the gastrocoele (G). In H (stage 12), the two areas of stain become confluent and remain so in later stages.

or pairs of forelimbs are developed. One develops from the normal somatopleuric mesoderm and the other from what would have gone into the formation of the splanchnopleure. These two sets of limbs, two rights and two lefts, are fully formed from each of the layers, but the splanchnopleure has not formed any of the smooth muscle which ordinarily would be present. There are visible, then, two definite defects associated with muscle formation, one in the region of the heart and one in the musculature of the gut. The stomodaeum is absent, as one would expect, and also the gills. The rest of the organism, however, is remarkably regular. These results change considerably our concept of how the muscle develops if we think conventionally of smooth musculature as being primitive, the cardiac as less primitive, and the somatic striate musculature as the most advanced. The facts seem to indicate a reverse of this sequence. The striate musculature is formed from the lateral mesoderm of the marginal zone. According to the findings of Vogt (1925, 1929), the areas of somite potency are well localized at the time when the marginal zone enters through the blastopore. It is at this time that it has its later characteristics impressed upon it.

The fact that the limb could be formed from splanchnopleure was already obvious if one considers the experiments of Harrison (1925) on the reversal of the mediolateral axis of the limb. Here, the limb was turned inside out, *i.e.*, the somatopleure and splanchnopleure were reversed with reference to the yolk. The splanchnopleure gave rise to perfectly good limbs and the somatopleure undoubtedly to good smooth musculature.

Nieuwkoop (*l.c.*), when he compressed the two sides of the material from which the yolk endoderm had been removed, showed cases in which he could distinguish the formation of a coelomic cavity which, however, had no splanchnopleure in association with it. It was entirely a somatopleuric cavity. He went also one step further and stretched the embryo after depriving it of its yolk upon a collodion membrane. In this case, development occurs in a manner strikingly reminiscent of a chick blastoderm, and again shows complete deficiency of all of the splanchnopleuric parts. No heart is formed except in those embryos in which the anterior parts had secondarily come together and even here the heart is markedly deficient.

The converse of these results is shown in Bacon's (1945) experiments. Bacon removed presumptive mesoderm from the outside of the embryo before the marginal zone had passed to the interior. When this marginal zone material was placed in contact with endoderm, it formed a heart rudiment. These heart rudiments, in some cases, were remarkably regular and were formed only after contact with the endoderm. There was no attempt to have the endoderm stay in relationship with the presumptive mesoderm, but they were dissociated after a period of contact, and yet remarkably normal hearts were secured, these hearts being kept entirely in isolation conditions. The fact that the tissue which ordinarily would have developed into normal striate muscle forms a heart after having been placed in contact with endoderm shows definite inductive action of the endoderm upon the marginal zone material to form heart. Since the endoderm itself was taken from the general mass of material and was not localized, it seems as though there is a generalized influence extending through the endoderm which acts upon presumptive mesoderm for heart formation. It would be interesting to check this thesis in connection with certain definite spatially specific parts of the endoderm. This, so far, has not been done. It would also be interesting to try various sections of the marginal zone to see whether they would respond in exactly the same way to the various parts of the endoderm. An interesting series of experiments could be evolved in which these two structures were tried each against the other in order to test out the inductive effects. This induction can occur at a stage earlier than nervous system induction, thus placing the Spemann organizer in an entirely secondary role with regard to the formation of the embryo as compared to endodermal inductions.

Holtfreter (1944) has been studying the actions and reactions of various parts of the yolk endoderm. In many cases, he has restricted his observations to the cells which are coming in through the ventral lip of the blastopore. He finds that these cells, which form distinctive elements, are polarized and that when they are isolated in salt solution one can get at certain stages a centrifugal spreading of the non-polarized cells, and a definite polarization with an amoeboid process in the polarized cells. They have a proximo-distal axis which seems, to some extent, to be controlled in normal development by the action of associated cells upon the endodermal mass. The amoeboid process is generally extending away from the central mass of cells and Holtfreter has been able to show that, if four or five cells are placed together, they tend to have their proximodistal axes arranged away from the center of the mass, the proximal portion being a rounded stub, while the distal part tends to be of the amoeboid type. Holtfreter likewise, in the same study, postulates that there is an independent movement in the cells of the endoderm, that this independent movement is responsible for a pulling of the structures so that they intrinsically occupy a definite mass and spread in a definite way. This, it seems, is endowing the endodermal cells with something a little more vital than one would expect, and probably more vital than the observational material would warrant. Whether the cells are pulling, or whether they have been displaced by the action of adjacent cells when acting in the normal organization of the embryo, is a question which still has to be decided. Holtfreter, however, has drawn our attention to an exceedingly important point in showing that, under conditions of isolation, the cells do certain definite things which seem to depend on mass aggregates and individual activity in response to that mass aggregate. No matter where the future of the endodermal situation will lead, these results will have a definite bearing upon our interpretation of how the floor of the gastrocoele is formed and what the effective forces are forming it. Certainly, the ventral lip of the blastopore has a much greater import in the organization of the entire venter of the embryo than we have hitherto supposed.

Kemp (1946) has changed various parts of the endoderm in relation to various other parts of the endoderm. He has removed sectors of the floor of the gastrocoele in the gastrular stage and has interchanged these either by cutting off one part of the embryo and then joining to it another embryo with a definite deficiency, or else by reversing, in some cases, the endodermal sectors in the lower portion of the neurula. These

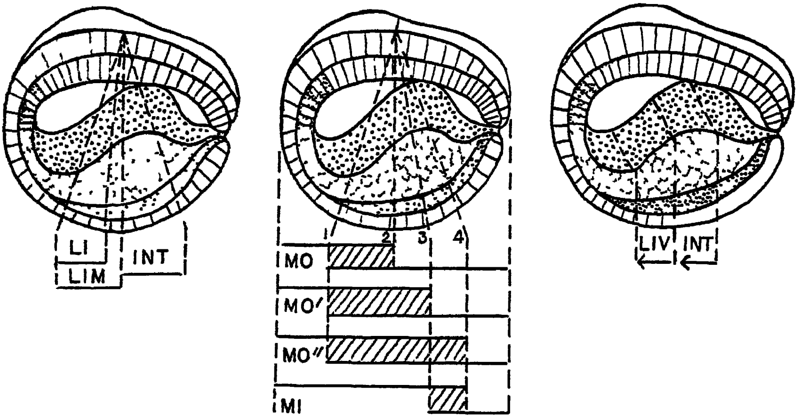


Figure 2. These three figures are redrawn from Kemp, 1946. The left figure shows the regions removed, comprising three series: Li, the anterior fourth of the mesenteron; Lim, the anterior third; and Int, middle third. The majority of the Li had functional tracts; Lim, non functional guts, all shortened, esophagus reduced or absent, liver and pancreas reduced, many heart defects; Int, ½ regulated completely, 92 per cent had functional tracts.

Middle figure—alteration of yolk endoderm arrangement by transection and reuniting of different levels resulting in duplication of parts of the mesenteron; MO, duplication of anterior third of mesenteron; MO′, anterior half; MO″, anterior two-thirds; MI, post segment. MO—80 per cent functional, 48 per cent completely regulated, enlarged or double pancreas, duplicate gall bladders. (Table 2—Kemp.)

Right figure—shows the regions rotated to reverse A–P axis of the segment. The effects in this series are most marked after the reversal of the Liv segment as indicated by the percentage of abnormality produced. His conclusion is that each endodermal anlage is histologically determined but is equipotential within itself.

Table 1*

Results of Duplication

Series	*Number in series*	*Per cent showing complete regulation*	*Per cent with functional tracts*	*Per cent with duplicated pancreas*	*Per cent with duplicated liver*
MO	25	48	80	20	0
MO′	35	5.5	23	62	6
MO″	32	0	25	100	6
MI	7	13	86	0	0

* From Kemp's Table 2, 1946.

have given rise to a set of rather interesting results, for in the course of this interchange he has been able to produce defects of the liver and of the heart. His results are interesting in showing that a certain regulation can take place at this stage in spite of the way in which the endodermal material has been combined.

As mentioned above, Holtfreter considers that the endodermal cells are polarized at the time they pass through the blastoporal lip and that from then on the endodermal cells in the floor of the archenteron are effective in stretching and accommodating themselves to the region of the floor, passing out and forward as well as laterally, and pulling material from the posterior region into position. They are non-polarized before they enter the blastoporal lip. Their behavior with other cells when isolated shows this polarization complex even though the other cells around them are not reacting, *i.e.*, the cells which normally would encompass the entire mass and make up the dorsal lip and its lateral projections, the marginal zone with its involution; all these can be separated from the action of the endodermal cells in the floor of the archenteron. Holtfreter goes one step further in saying that the surface coat material which is pulled in with the ventral lip as well as with the dorsal lip acts in forming the linings of the cavities of the body as they are later found. This would seem to place a little too much emphasis upon the surface coat, which, although it comes in with the material, is strikingly reduced in the lower layer as compared with its amount on the outside of the egg itself.

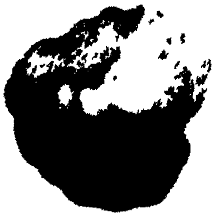
Holtfreter (1933) gave one of the best examples of what can happen in his exo-gastrulation studies. In this case, after having treated the embryos with hypertonic solutions, he secured complete exo-gastrulation with the mesoderm forming its parts lying on top of the endoderm and forming gill arches and body form without the action of the ectoderm. This is an interesting check on the results which Nieuwkoop obtained, for here we have the situation exactly in reverse. The Spemann organizer has not acted in forming the nervous system. Instead, however, during exo-gastrulation, the mesodermal parts have been acted upon, forming a perfectly good head region without, of course, the nervous system, forming gill arches, and forming the body which has the general conformation of what it would have had at the same stage if the embryo had developed with the ectodermal covering around it. The ectoderm, potentially capable of making nervous system as well as body covering, has formed only a bladder of material in which there is no differentiation except that of a common epidermis.

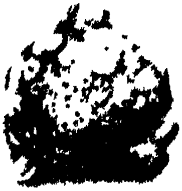
In his 1938 and 1939 studies, Holtfreter has performed other experiments showing still another part of the reaction of the endoderm. When ectoderm and mesoderm are united, a vesicle is formed containing connective tissue, but little else. When the endoderm is added to the isolated bladder of material, there is definite structural conformation and arrangement of the mesoderm, the formation of certain muscular parts in connection with the endoderm, the whole being surrounded by ectoderm. The ectoderm in this situation does not influence the mesoderm nearly so much as does the endoderm, for in the endodermal parts we find a definite uniting of structure, with the mesoderm contributing a muscle layer and the outside sheet of the gut which is differentiating from the endoderm. Here, the potentiality of the gut region is clearly marked; it has a much greater definitive potency for calling forth from mesoderm a differentiation unlike the connective tissue which would develop within the ectodermal bladder by itself. This is another demonstration of the effective facility with which the endoderm acts upon associated structures.

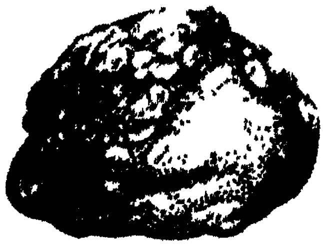
The experiments of Nieuwkoop and Holtfreter strikingly demonstrate that, in the gut endoderm, we have certain capacities for organization which have been neglected in looking upon the embryo as a whole. Its capacity to organize by induction, the facts that the heart is missing without the influence of the yolk endoderm and that the gills are completely missing, had been noted in Stöhr's (1931) experiments. He suspected that the endoderm exerted some influence upon the reacting structures in causing the formation of heart. He did not, however, suppose that the endoderm was directly responsible for smooth musculature of the gut and that the somatopleure and splanchnopleure together could form interchangeable structures.

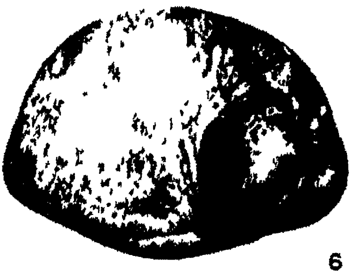
It is interesting, therefore, to review the morphogenesis of the muscular system, for the splanchnic mesoderm, if not in contact with endoderm, will form striate muscle under the influence of the somatopleure. Any part of the marginal zone which is potent to form muscle striate in character can be induced, according to Bacon's results, to form heart in the presence of endoderm. The reversal of the limb bud (placing the splanchnopleure on the outside, the somatopleure on the inside) gives rise to a perfect limb. All these things tend to show that there is a dominant action exerted by the endoderm which has a part in the induction of muscular structure. It also shows that the final determination of such structure lies not only in its general position, but also in its definitive relationship to other tissue influences.

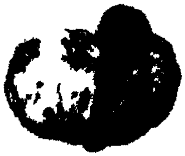
In the light of Holtfreter's (1939) experiments, it was thought advisable to carry the isolation of the yolk back further into the foregrounds of embryonic history than he had done. Yolk endoderm of the late blastula, when grown in culture (Nicholas, 1945), assumes about the same form relationship as would be found in the normal animal. When isolation experiments were performed upon young blastulae, it was found that not enough of the plasmagel of the external coat material was present in the endoderm to permit its attainment of form. The experiments in which ectoderm is combined with yolk endoderm, as shown in FIGURES 3 to 8, are more successful when different proportional amounts of ectodermal-endodermal material were used. When yolk endoderm material is taken from the early blastula and cultivated without an external cover, the results are rather clearly shown (FIGURE 4). The cells aggregate and so does the nuclear and pigment material; there is a fine granular


3


4


5


6


7

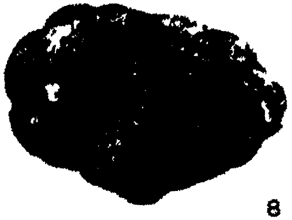

8

FIGURE 3. Isolation study of yolk endoderm from early blastula enclosed in envelope of ectoderm. Cultured 5 days. A thin coat covers all of the exposed endodermal material. The ratio of endodermal material to ectodermal cover is 2 to 3.

FIGURE 4. Yolk endoderm of late blastula cultured in the absence of ectoderm in double Holtfreter's solution. Note the cell aggregate and the clumping of pigment and nucleus within the constituent cells. Cultured 4 days. There is no ectoderm present and fine granular material was being exuded from the loose cells located in upper part of the diagram.

FIGURE 5. Yolk endoderm isolate, mid-blastula, in combination with a small ectodermal piece, shown to the left of the figure. The ratio of endoderm to ectoderm is 5 to 1. After 5 days of culture, disintegrative action began, as shown by the loose cells in the upper left of the figure.

FIGURE 6. Yolk endoderm isolated in early blastula and combined with ectoderm in 4-to-1 ratio. There is here, as in the previous figure, a distinct antero-posterior lengthening of the yolk endoderm. There is a rolling of the cells at the inferior margin of the figure, showing that movement of the endodermal constituents continued under these conditions.

FIGURE 7. A 1-to-1 combination of yolk endoderm of an early blastular fragment with ectoderm from the blastular cap. After 7 days of culture, there is no axial or other differentiation, probably due to the small size of the piece.

FIGURE 8. Isolate consisting of 1 part of yolk endoderm combined with 5 parts of ectoderm. The ectoderm contains a small ball of endodermal material which in 6 days of culture did not differentiate.

type of exudate, probably made up of the yolk platelet content, and in spite of the fact that the isolates were kept in double Holtfreter solution, which is osmotically favorable to endoderm preservation, they disintegrate usually within 24 hours. If, however, ectoderm is combined with the yolk endoderm, as shown in FIGURE 3, there is a tendency for the endoderm to receive something from the ectoderm which acts as an external cover preventing the disintegration of both the ectoderm and the endoderm.

The same relationship is shown in FIGURES 5 and 6, which are examples of about the minimum of ectoderm combination which is necessary for the development of the yolk endoderm taken from the early blastula stages. FIGURE 5 shows rather interestingly that the external cover over the yolk endoderm is not adequate to prevent sloughage, which can be seen in the large, loose, white mass of cells to the left of the figure. An isolate (FIGURE 6) in which there was about 25 per cent of ectoderm at the time of its first isolation, has a complete coverage and can be maintained for a considerable period of time. The elongation of the yolk endoderm under these conditions simulates the normal axial relationship. The ectodermal covering is found near the anterior end of one of these masses and near the posterior end of the other, as judged by their general morphological appearance. Ectoderm by itself does not seem to have a definite effect upon the polarization of the yolk mass. An early isolate in about the 128-cell stage is illustrated. Here, an increasing amount of ectodermal material is necessary if the isolate is to survive in double Holtfreter solution. As in previous figures, one sees the relationship of the ectoderm to the yolk endoderm in that there is a consistent smooth covering of material on the outside of the endodermal cells which they themelves do not produce in amounts adequate to cover their surface. In FIGURE 8, also an isolate, in which much more ectoderm was taken and where the relative amount of the endoderm was small, the ectoderm completely covers the endoderm, the endoderm remaining a small, nodular mass inside the ectoderm. While it retains its normal components, there is no indication that there is any development of polarity of axial determination.

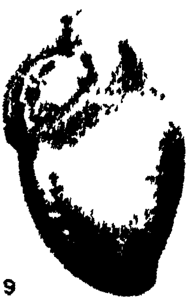
In this series, then, it is shown that material from the early blastula lacks the capacity for the formation of a superficial plasmagel in sufficient quantity to prevent content losses to the surrounding medium (in this case a double Holtfreter solution).

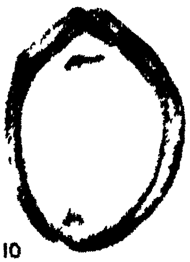
When varying proportions of ectoderm, taken from any region of the blastula, are included with the early blastular endoderm in the isolation, the endoderm can be maintained without disintegration. The self-differentiating powers of the yolk endoderm cannot be demonstrated unless it secures additional material from potential ectoderm cells which increase its surface membrane and enable the endoderm to undergo its own arrangement.

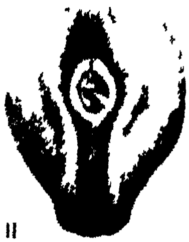
This is not regarded as a direct morphogenetic effect upon the endoderm by the ectoderm, but as a mechanical factor supplied by this material in the form of plasmagel. Elongation and the expression of axial tendencies are obtained only when the yolk endoderm of the early blastula is present in adequate proportions. A four-endoderm-to-one-ectoderm combination, as shown in FIGURE 6, is optimal. When the proportions are reversed, as in FIGURE 8, one endoderm to four ectoderms, the endoderm shows no elongation but remains as a rounded ball with an ectodermal cover.

A separate series of observations was made upon early blastulae in which the roof of the blastocoele was removed, and a piece of the transparent vitelline membrane was cut to fit the opening and inserted so that the ectoderm was fitted around the margins of the cover. It was possible, by this method, to observe the formation of the floor of the blastocoele and the way in which the various parts came into their organization complex during the course of development. The vitelline membrane remains transparent, is retained by the organism without apparent damage to the adjacent tissues, and, since it softens during development, serves to separate the various parts during the critical period in the organization of the floor of the archenteron. It was noted early by this observational method that there was a pattern of cells arising near the midpoint of the blastula in the floor of the blastocoele. There are generally between 16 and 28 of these rather large cells, which appear in the form of the mosaic in the floor of the blastocoele. These are the cells which have formed in the pathway of ingression (see FIGURE 1). Whether they are actually cells that were originally on the outside cannot be definitely stated. There are probably many more cells, not distinctive in either their form or their position, which come in at this time, and certainly the 16 to 24 which projected the pebbled surface on the floor of the blastocoele can be localized as being morphologically in the chain of ingressive material. They later are found under the stomodaeal pit as the blastocoele becomes reduced.

In addition to the experiments reported above, in which the observations of the cells in the floor of the blastocoele could be made, additional experiments were conducted in which the entire cover of the early gastrula was removed. In this series of experiments, the neural folds developed at the marginal limits of the denuded area; the yolk endoderm was exposed, and the results of the previous series of experiments were checked. Just anterior to the midpoint of the blastocoele floor, the pebbled series arises, forming a portion of the floor of the archenteron. Above this develops the stomodaeal pit, while posteriorly the proctodeal pit is found. These two structures seem to develop entirely from the yolk endoderm, which is specific for them. The factors which have to do with the formation of the stomodaeal pit and the proctodeal pit certainly have their origin intrinsically in the yolk endoderm. The occurrence of the lateral neural folds presented somewhat of a problem, since this material was underlain only with marginal zone material which normally would have become lateral mesoderm. When the substrate is marginal zone material, however, and not the usual chorda mesoderm, it has the capacity and competence for forming ectoderm into nervous tissue (see FIGURE 10). The neural folds, while poorly developed anteriorly,







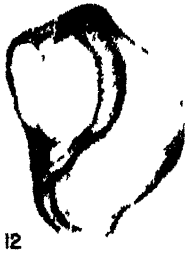


FIGURE 9. The utilization of the transparent vitelline membrane as a blister through which observations upon the formation of the stomodaeal invagination are made. The neural folds are held apart in the region of the mid- and hindbrain.

FIGURE 10. A culture of the yolk endoderm held in place by the ventral ectoderm and the marginal zone. The operation removed practically the upper half of the blastula. The formation of the proctodaeal and stomodaeal invaginations is regular and occurs because of reactions intrinsic to the ventral half of the early blastula.

FIGURE 11. A small cap was removed from the dorsal part of the blastula and observations made upon the midpoint of the blastocoele floor (*cf.* FIGURE 1). The cellular mosaic developed regularly in the floor and is later, by the involution of the blastocoele, converted into the floor of the archenteron beneath the stomodaeal groove.

FIGURE 12. Culture of early blastula with same type of operation as shown in FIGURE 11 but without vitelline cover. The growth processes in the endoderm are less regular and a part of the endoderm has been displaced in the healing process.

show good structural relationships, and the sections through these individuals show perfectly definite systems.

If the experiment is performed later, only part of the roof of the mid-gastrula is removed and the cellular reactions which are going on in the floor of the gastrocoele can be observed. The nervous system, instead of consisting of half the nervous system in the regions which have been separated by the removal of the roof of the blastula, now is practically complete.

There is a distinct difference, then, in the nervous system effect of the two operations, for in the first case the nervous system forms really as a vestigial type of regenerate; in the second case, where only a small portion of the roof of the blastula was removed, it is practically complete.

Discussion

The question immediately before us to be considered is not that of the localization of the parts of the gut in the yolk endoderm. This is adequately treated in Kemp's (1946) paper. In general, the results given here support Holtfreter's (1939) conclusion that there is early localization of the gut. The stomodaeum and proctodeum are determined in the late blastula, and there are indications that the morphological components which determine this change are the elements which have reached the inside of the blastocoele by ingression. This is a rather important point, for it shows an early localization of gut structure as well as an antero-posterior localization of the materials. The parts, however, between the stomodaeum and the proctodeum seem to have only a general relationship.

The general relationships of the yolk endoderm are more important for the present discussion than the absolute localization of parts such as liver, or the parts of the liver. As has been pointed out previously, the yolk endoderm has a generalized capacity which can call forth from the surrounding tissues certain definite elements. This is also clear from Nieuwkoop's experiments, for, in his study of the origin of the germ cells, one can see the various ancillary actions which the yolk endoderm exerts. It is positive in the formation of the stomodaeum and the heart which shows from Stöhr's experiments (which are the converse of Bacon's), as well as in various experiments having to do with the gills and their formation. The endoderm, then, is exceedingly potent in the formation of structures at the anterior and ventral aspects of the embryo. The fact that splanchnopleure does not develop if the gut is absent shows a definite directive effect of the yolk endoderm upon the histogenesis of what might be regarded as a generalized muscular structure. The inhibiting action of the yolk endoderm which prevents the splanchnopleure from forming limbs is another important indication of its action.

Yolk endoderm is probably the largest source of the chemical materials which will be transformed by the organism. These are the rough products from which the later chemical materials are going to be formed. It is possible that the yolk platelets which have been separated into the ectodermal cells and which are part of their composition are changed in their values with relationship to the chemical constituents by the cells themselves. If this is so, we have a rather clear indication that certain of the materials which are found predominantly in the nervous system later are developing from the raw materials of the yolk endoderm which have been carried about and placed in a definite location by the cellular activity itself. The greater reservoir of these materials still remains in the yolk endoderm which is left behind.

In watching the development of the amphibian forms, one is frequently impressed with the fact that the material found within the gut, consisting of the yolk endoderm particles, which have been enclosed by the gut material, does not completely digest, and, while there is an extraction of material from the yolk, there is still sufficient residue left to make quite a sizable collection of material. This material is relatively inert and should be investigated in order to get the difference between the chemical substances which are involved in the yolk platelet formation and the final residue after its extraction by the organism. This should form a field of rather useful investigation for the chemical understanding of what has gone into the organic relationships which have been formed by the embryo itself.

The yolk endoderm, then, has been shown in the present discussion to be responsible for the formation of the stomodaeum, to take part and be active in the formation of gills, and to possess the capacity for induction of the heart from indifferent mesoderm or from mesoderm which is prospectively significant for the formation of other structures than the heart. Since the yolk endoderm is so active in the formation of anterior structures, it should be looked upon as a possible factor influencing the formation of both the hypophysis and the thyroid. It has a corollary in Nieuwkoop's findings that the endoderm is responsible for the induction of germ cells from the lateral plate. From Nieuwkoop's work, it is clearly shown that the muscle developing from the somites in the splanchnopleuric region will not develop as the smooth muscle of the gut, but acts in the absence of the yolk endoderm as somatopleure which has the capacity for developing limbs at the normal regional level similar to that found in the somatopleure. When the yolk endoderm is removed in early stages, the heart is atypical or absent. When the yolk endoderm is cultured in mid-blastula stages, it elongates, possesses a definite plasmagel coat, and develops the invaginations normal to the formation of a stomodaeum and proctodeum. In earlier stages of the blastula, some material from the outside of the egg must be incorporated along with the yolk endoderm in order to maintain it as a unit and have it continue its development. One can secure the development of practically normal form and extension if the yolk endoderm is 25 per cent of the quantity of the graft and if it is covered by external material. This points to the necessity for the normal operation on the yolk endoderm of material from the outside of the egg. The fact that the mid-blastula stages can be cultured shows that sufficient material of a nature similar to that of the external coat of Holtfreter is present in the cells of the yolk endoderm and can take care of its relationship with the external environment.

The development of the chemical substances and the enzymes which are effective in energy transformation may occur in the cells that carry a maximal or minimal amount of the yolk platelets. These are particularly rich in the endodermal cells, as is shown by the cytochrome oxidase and probably by the adenyl-pyrophosphotase development.

Bibliography

Bacon, R. L. 1945. Self-differentiation and induction in the heart of *Amblystoma.* J. Exp. Zool. 98: 87–125.

Child, C. M. 1941. Patterns and Problems of Development. University of Chicago Press. Chicago.

Daniel, J. F., & E. A. Yarwood. 1939. The early embryology of *Triturus torosus.* Univ. Calif. Publ. Zool. 43: 321-356.

Hamburger, B. 1947. *Experimental Embryology:* 973–980. Encyclopedia Britannica.

Harrison, R. G. 1925. The effect of reversing the medio-lateral or transverse axis of the fore-limb bud in the salamander embryo (*Amblystoma punctatum* Linn.). Arch. Entw.-mech. 106: 469-502.

Holtfreter, J. 1933. Organisierungsstufen nach regionaler Kombination von Entomesoderm mit Ektoderm. Biol. Zentralbl. 53: 404–431.

1936. Differenzierungspotenzen isolierter Teile der Urodelengastrula. Arch. Entw.-mech. 138: 522–738.

1939. Gewebeaffinität, ein Mittel der embryonalen Formbildung. Arch. exp. Zellforsch. 23: 169–209.

1944. A study of the mechanics of gastrulation. Part II. J. Exp. Zool. 95: 171–212.

Kemp, N. E. 1946. Regulation in the endoderm of the tree frog *Hyla regilla.* Univ. Calif. Publ. Zool. 51: 159–184.

Nicholas, J. S. 1945. Blastulation, its role in pregastrular organization in *Amblystoma punctatum.* J. Exp. Zool. 100: 265–299.

Nieuwkoop, P. D. 1946. Experimental investigations on the origin and determination of the germ cells, and on the development of the lateral plates and germ ridges in Urodeles. Arch. Néerl. Zool. 8: 1–205.

Rhumbler, L. 1902. Zur Mechanik des Gastrulationsvorganges, insbesondere der Invagination. Eine entwicklungsmechanische Studie. Arch. Entw.-mech. 14: 401–476.

Roux, W. 1885. Beiträge zur Entwicklungsmechanik des Embryo. III. Über die Bestimmung der Hauptrichtungen des Froschembryo im Ei und über die erste Theilung des Froscheies. Abhandl. Entw.-mech. 2: 277–343.

Schechtman, A. M. 1934. Unipolar ingression in *Triturus torosus*: a hitherto undescribed movement in the pregastrular stages of a Urodele. Univ. Calif. Publ. 39: 303–310.

1935. Mechanism of ingression in the egg of *Triturus torosus.* Proc. Soc. Exp. Biol. & Med. 32: 1072–1073.

1937. Localized cortical growth as the immediate cause of cell division. Science 35: 222–223.

Stöhr, P., Jr. 1931. Beobactungen zur Organentwicklung bei erythrozytenfreien Amphibienlarven. Arch. Entw.-mech. 124: 705–746.

Vogt, W. 1925. Gestaltungsanalyse am Amphibienkeim mit örtlicher Vitalfärbung. Vorwort über Wege und Ziele. I. Teil: Methodik und Wirkungsweise der örtlichen Vitalfärbung mit Agar als Farbträger. Arch. Entw.-mech. 106: 542–610.

1929. Gestaltungsanalyse am Amphibienkeim mit örtlicher Vitalfärbung. II. Gastrulation und Mesodermbildung bei Urodelen und Anuren. Arch. Entw.-mech. 120: 384–706.

THE ROLE OF NERVES IN AMPHIBIAN LIMB REGENERATION*

By S. MERYL ROSE

Smith College, Northampton, Massachusetts

THE first study of the influence of nerves on regeneration was reported by Tweedy John Todd in 1823.[60] He had discovered that division of the sciatic nerve in the limb of a salamander prevents regeneration distal to the level of division of the nerve. He further stated that division of the nerve causes a regenerate already developing to shrivel and waste. These observations lay unused for many years, until interest revived and Todd's paper was rediscovered early in this century. Several papers appeared soon after the year 1900, some affirming and others expressing doubt[9, 37, 65] that nerves are necessary during regeneration. The question was finally settled in the affirmative by Wolff in 1910[66] and by Walter in 1912.[64] Later, Schotté made it clear that if a nerve supply to a salamander's limb is interrupted, regeneration does not begin until after nerves have regenerated and re-entered the region adjacent to the level of amputation.[44] The expressions of doubt had apparently been due to faulty operations in some cases and to failure to recognize regenerated unmyelinated fibers in others.

Prior to the analysis of the role of nerves in regeneration, attempts were made to isolate the various components, sensory, motor and sympathetic, and to determine thereby which are necessary for regeneration. At first, it seemed that only the sensory supply was necessary because regeneration could occur in the absence of the section of the spinal cord supplying the limb plexus if the dorsal root ganglia and the sensory fibers were not severed.[64] This was regeneration in the absence of a motor supply. Schotté reinvestigated the question of which nerve component, motor or sensory, is needed and found inconsistencies in his results. For example, sometimes the sensory supply alone seemed to be enough for regeneration, at other times a regenerated motor supply was sufficient, and even the original motor supply seemed to be enough in a few cases.[41] It was suspected that the inconsistent failures in regeneration were caused by chance interruption of the sympathetic nerve fibers.[42] Following this hypothesis, attempts were made to eliminate the sympathetic supply in some cases, and in others to remove everything but the sympathetic. The results were never completely consistent, but numbers favored the sympathetic nerves as the important component for limb regeneration.[43]

This conclusion was widely accepted, but it did not go wholly unchallenged. Locatelli pointed out that the sympathetic nerves, especially the communicating rami, are extremely small, and the rami, at least, cannot be seen through the dissecting microscope.[20] She refused to accept the conclusion that the sympathetic component is important in regeneration, because of the extreme difficulty of the operations and the failure to confirm them histologically. Her own experiments reaffirmed that limbs can regenerate with sensory fibers alone. Working with the nerves supplying the hind limbs, she dissected around the roots of nerves 17 and 18 in order to remove communicating rami and then pulled 17 and 18 free from the cord. The operation was done in such a way that both ventral and dorsal roots were torn, but the dorsal roots were torn proximal to their ganglia. This left the sensory fibers of the limb still attached to their cell bodies but severed the connections between motor fibers and their cell bodies in the cord. With only the sensory fibers intact, regeneration was normal. In some cases, the sensory component of the 18th nerve by itself supported normal regeneration, but there was no regeneration if the 18th ganglion was extirpated. These results were taken to mean that the 18th dorsal root ganglion had a regeneration-promoting quality peculiar to itself.[20]

---

* Contributions from the Department of Zoology, Smith College, No. 215.

This was the state of the problem in 1929. Doubt had been cast on the claims for the sympathetic, and only part of the sensory supply seemed necessary for limb regeneration. In 1942, Singer began to publish an experimental re-analysis of the problem. His first finding was that post-ganglionic sympathetic fibers passing to the anterior limb in *Triturus* do not join the mixed nerves of the brachial plexus by way of communicating rami. Instead, the post-ganglionic fibers leave the sympathetic cord and constitute a separate nerve, the subclavian, closely applied to the subclavian artery.[52] Older observations (see Singer, 1942a, for bibliography) show a subclavian nerve to be the usual pathway for sympathetic fibers leading to an anterior limb in Urodeles. This fact invalidates the earlier experiments on the sympathetics. A reinvestigation showed that complete unilateral sympathectomy in the anterior region did not prevent normal regeneration on the side concerned.[53] In addition, there are reports that forelimbs had regenerated normally after their subclavian arteries had been severed.[14, 44, 67] Since it would be extremely difficult to sever the subclavian artery without also severing the nerve, these observations lend weight to the conclusion that limbs may regenerate without sympathetic nerves.

Further careful re-analysis by Singer, with operations checked histologically, show the truth of the thesis that the sensory is the only component which by itself can support normal regeneration.[53, 54, 55] However, this does not mean that sensory neurones have a peculiar quality enabling them, and them alone, to support regeneration. This is not true, as Singer shows, for, although a normal motor supply will not support regeneration,[55] a regenerated motor supply will.[56] The key to an understanding of this paradox comes from a consideration of the number of fibers involved. There are many more motor fibers after regeneration from a ventral root because the axones branch repeatedly as they regenerate. It has been determined that the number of fibers necessary for limb regeneration in *Triturus* is in the range of one-third to one-half of the total number.[57] The normal motor component does not contain a third of the total number of fibers, but there are enough fibers in a regenerated motor supply to satisfy the threshold requirements. The sensory component is the only normal component with enough fibers to promote regeneration by itself. In fact, there are so many sensory fibers that not all of them are needed. Unless the assumption is made that the regenerated motor fibers have acquired a new quality not present in the normal motor supply, it seems likely that neurones of all components possess the quality necessary for regeneration. Certainly, sensory and regenerated motor nerves do.

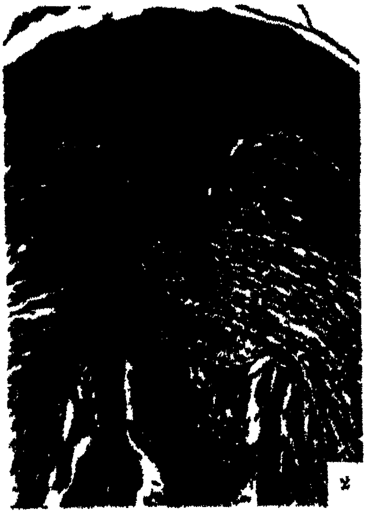
In the course of her experiments, Locatelli discovered that limb nerves deviated from their normal course and, when made to end near the base of a limb, stimulated the development of a new limb over the end of the nerve.[19] The idea that nerves possess specific morphogenetic potency was suggested, but had to be discarded after a demonstration that limb nerves deviated to adjacent territories stimulated regeneration of structures whose form was determined by the territory rather than by the nerves.[11, 2] For example, when the limb nerve ended at the base of the dorsal crest, a piece of dorsal crest developed.

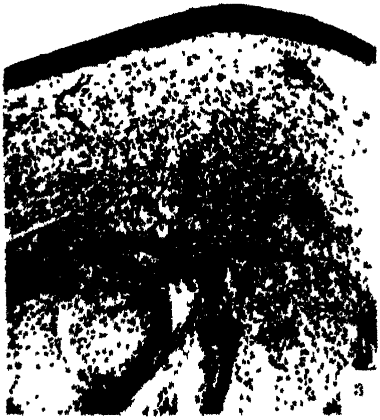
We now face the problem: How do nerves stimulate regeneration? In trying to discover their function, we are aided by the knowledge that embryonic limb buds can grow and differentiate without a nerve supply, as first demonstrated by Harrison.[13, 12] Since nerves are not necessary in embryonic development, it would seem that our search should begin with that phase of regeneration which differs from embryonic development. That phase is the first phase in regeneration, really a preparatory one, during which the old tissue organization at the level of amputation is lost, and free cells, apparently released from old tissues, collect and form a bud.[3, 45, 21, 62, 31, 28, 6, 59] This change from organization as tissues to a cellular type of organization is known as dedifferentiation. The bud or blastema, once it is established, resembles an embryonic limb bud, and from that time on its development is similar to the development of an embryonic limb. It is here, during the preparation of the blastema, that we shall first look for an influence of nerves.

The fact that nerves are required during this early phase of regeneration is clear: a blastema does not form in the absence of nerves. Only lately has there been much work on the causes of failure to regenerate after denervation. Schotté[44] and, later, Butler and Schotté[5] and their coworkers[50, 51] demonstrated that denervated limbs of young, larval salamanders begin to dedifferentiate in the region of an amputation surface but fail to regenerate because dedifferentiation and resorption continue unchecked until whole limbs disappear. The reason for failure to develop a blastema is quite different in older limbs and it is with them that we shall begin our analysis.

It is necessary to digress for a moment in order to introduce a regenerative process which we shall show to be influenced by nerves. It has been known for a long time that there is no regeneration from a wound sealed with skin.[61, 58, 8] This is not the result of a physical block to blastema growth. What a complete skin seal does is this: By covering a wound, it prevents an epidermal epithelium from spreading over and into the wound, thus barring close contact of epidermal wound epithelium with underlying tissues. In normal regeneration, it is the close contact of wound epithelium with underlying tissues which causes histolysis of those internal tissues and the release of free cells as shown by Jeffimoff[15, 16] and by Polejaiev.[23] A complete skin over a wound prevents regeneration, not by preventing growth of a blastema—far from it—but because it prevents dedifferentiation which provides the cells for the blastema. If a limb wound is covered in such a way that one corner of a skin seal is free, a wound epithelium can migrate in under the skin covering the wound and do its work.[15] Intense internal histolysis leading to blastema formation results and, instead of the skin acting as a physical barrier to growth, it, too, undergoes histolysis.[8] Not all epidermal wound coverings, even in salamanders, cause extensive dedifferentiation of the tissues underlying them. For example, although epithelia from limb, tail or abdomen will support limb regeneration, wound epithelia provided by skin of the head or back will not.[16, 29]

Passing, now, to a non-regenerating animal, the adult frog, we find that one of the causes for failure to regenerate is a change in the epidermis during metamorphosis.[24, 49] Miss Gidge and I obtained good regeneration of forearms and wrists in adult frogs after substituting tadpole skin for their own adult skin.[7] A wound epithelium provided by the transplanted tadpole skin migrated over and into the wound of amputation and initiated extensive dedifferentiation. This is very limited in adult frogs if their own epidermis contributes the wound epithelium.[36] The heavy dermis of adult frogs may also be a factor in preventing regeneration because it, along with epidermis, soon closes over a wound, forming a barrier between epidermis and underlying tissues.[33] However, adult anuran limb epithelium can be stimulated by treatments with NaCl[32, 33, 35] or with other irritants[26, 27] and will then act as a younger epithelium and, in its turn, cause dedifferentiation even of dermis with the concurrent release of cells to be used in a blastema. FIGURE 1 is a section of a salt-treated frog limb. If it is compared with FIGURE 2, a control limb of the same age, it will be seen that dedifferentiation had just started in the salted limb. Bone, for example, had begun to erode away. Comparison of a salt-treated limb (FIGURE 3) and a control limb (FIGURE 4) a few days later shows intense dedifferentiation and concurrent blastema formation in the salt-treated limb while the control has already





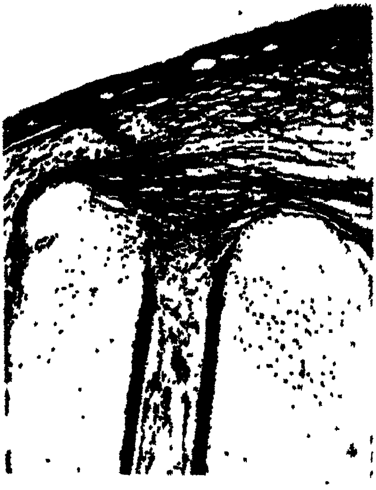


FIGURE 1. A longitudinal section through the tip of a salt-treated frog's limb 19 days after amputation and 10 days after the last salt treatment. Internal dedifferentiation is just beginning. Erosion of bony matrix and the presence of several large, darkly stained, free-fused osteocytes may be noted. X70.

FIGURE 2. A longitudinal section through the tip of an untreated frog's limb 19 days after amputation. This shows no active dedifferentiation and is to be contrasted with the salt-treated limb in FIGURE 1. X70.

FIGURE 3. A longitudinal section through the tip of a salt-treated frog's limb 27 days after amputation and 18 days after the last treatment. A large blastema of mesenchymal cells has formed, beneath which is a zone of dedifferentiating old bone and new cartilage. X45.

FIGURE 4. A longitudinal section through the tip of an untreated frog's limb 27 days after amputation. Already, fibrous scar tissue and cartilage, flanking the bone, have completely differentiated. This premature differentiation is in marked contrast to the undifferentiated state seen in FIGURE 3. X45.

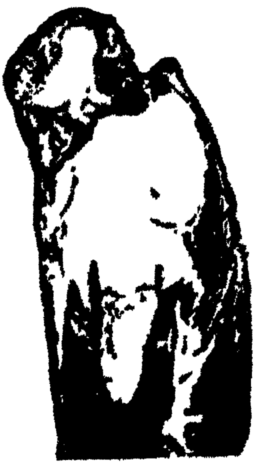
healed with a cartilaginous callus surrounded by fibrous scar tissue. The dedifferentiation had apparently been initiated by a stimulated epithelium.

The best of the regenerates obtained from adult frogs by either salt treatments or tadpole skin transplantation, after amputation through the forearm, had normal new forearms and wrists but were abnormal in the more distal regions (FIGURE 5). We shall consider the cause of this abnormality later. At present, we are interested in salt treatments only because of their action in stimulating a wound epithelium.

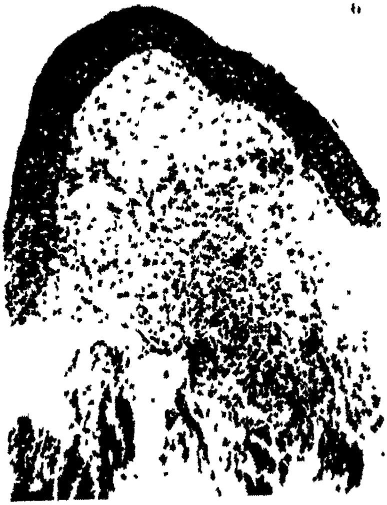
Returning from the digression with the knowledge that an active wound epithelium causes internal histolysis, we can now show that nerves are normally involved in the activation of the wound epithelium in adult salamanders, as Miss Oliver and I[22] have recently demonstrated. FIGURE 6 shows a blastema on a normal limb of the newt, *Triturus viridescens.* Contrasted with this is the tip of a non-regenerating denervated stump of approximately the same age (FIGURE 7). In the denervated limb, there has been very little dedifferentiation. Only the cut muscle fibers have degenerated. Bone, a good landmark, remains intact to the end of the stump. This stump failed to develop a blastema in the absence of nerves because it failed to dedifferentiate. Notice the wound epithelium, only a few cells thick and composed almost entirely of squamous cells. Now compare a completely denervated limb of the same age which had received baths in a strong NaCl solution (FIGURE 8). The wound epithelium is much thicker and extends down into the wound. This time, there is active dedifferentiation. Bony matrix is being eroded away and, in every case, multi-nucleate, fused osteocytes lie in the wake of the process. The conclusion is drawn that nerves promote dedifferentiation in the adult limb stump.

Dedifferentiation is not the only process for which nerves are needed during regeneration, as we learn from consideration of the ultimate fate of the limbs in question. Cells in the blastemata of normal limbs grow rapidly and remain undifferentiated until a large group of cells has formed. Only then does differentiation begin. In the denervated, salt-treated limbs, there is appreciable dedifferentiation, but the most striking phenomenon is the rapid differentiation of new tissue, cartilage, and fibrous scar tissue, with a concurrent failure of growth. In the untreated, denervated limbs, few cells dedifferentiate, but the few which do so almost immediately redifferentiate again into scar tissue and cartilage. In FIGURE 9, showing a denervated limb of the same age as those in FIGURES 7 and 8, premature differentiation of cartilage along the edge of the partially eroded bone may be seen. This is never observed in salamander limbs with a normal nerve supply. It would seem that nerves promote growth without differentiation, or, rather, shift the balance between growth and differentiation toward growth. In this way, a sufficiently large mass of tissue is available to form a limb by the time differentiation begins.

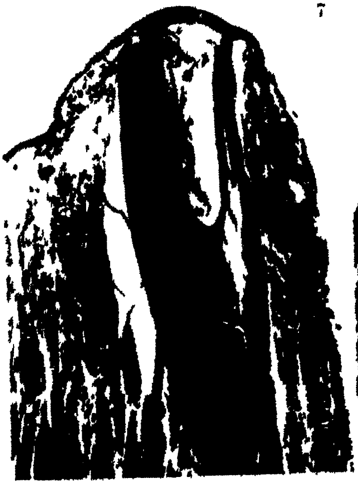
5



6



7



8

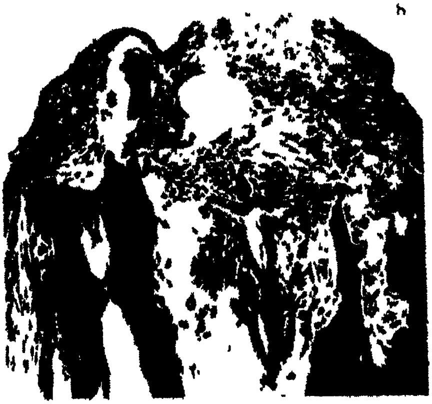


FIGURE 5. One of the best of the frog limb regenerates obtained after transplantation of tadpole skin. The limb had been amputated through the middle of the forearm. Pictured here is the regenerated distal forearm and normal functional wrist. It is to be noted that in this and in all of the induced regenerates the distal part of the hand and the fingers are either abnormal or missing. X12.

FIGURE 6. A cone-stage blastema on the forearm of an adult *Triturus* 23 days after amputation. Dedifferentiation had progressed as far as the tissues at the base of the photograph, and rapid growth of the blastema had begun. X70.

FIGURE 7. The tip of an adult *Triturus* forearm 23 days after complete denervation and amputation. There has been very little loss of structure. Cut muscle fibers have degenerated, but bone is still intact even at the level of amputation. X70.

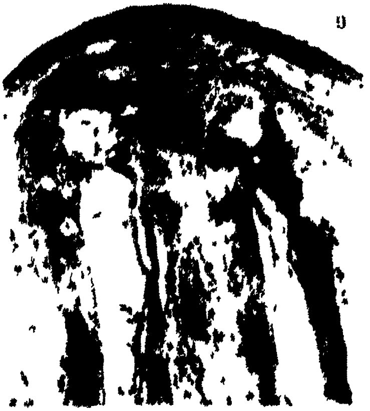
FIGURE 8. The tip of a salt-treated adult *Triturus* forearm 23 days after complete denervation and amputation. The limb had been treated with strong NaCl solutions for the first 14 days after the operation. The greatly thickened wound epithelium, typical for treated limbs, is to be contrasted with the thinner wound epithelium on the untreated limb in FIGURE. 7. Dedifferentiation of bone had begun in the treated limb. X70.

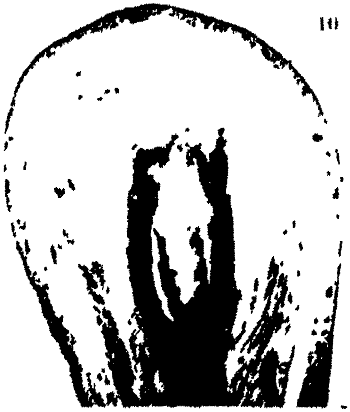
There is supporting evidence for this belief. If a denervated newt limb is re-amputated after 23 days, some nerve fibers have apparently reentered the limb and a small fraction of these fresh stumps will regenerate limbs, but most of them are insufficiently innervated. These produce callus and scar tissue, as seen in FIGURE 10. The same type of growth was observed by Walter many years ago[64]; it was also caused by insufficent innervation. There can be little doubt that nerves are important for limb regeneration not only because they favor dedifferentiation but also because they promote rapid growth of the dedifferentiated cells. If growth fails and differentiation begins prematurely, all the cells in the vicinity of bone are transformed to chondrocytes and others closer to the periphery become fibroblasts.

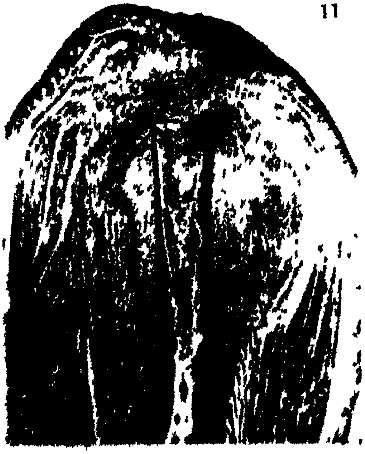
This combination, inadequate dedifferentiation followed by poor growth and rapid dedifferentiation ending with callus formation, is found not only in denervated and x-rayed limbs, but also in the limbs of naturally non-regenerating vertebrates.[34, 35] FIGURE 11 is a section through a 9-day stump of a frog's limb. Already, its limited dedifferentiation and premature differentiation have resulted in the formation of callus and scar. The same picture is seen in a lizard limb stump (FIGURE 12). Here, too, fiber and cartilage form prematurely, sooner than these same tissues form in the regenerating tail of a lizard, as shown by Barber.[1]

The question naturally arises: Do these limbs fail to regenerate because of insufficient innervation? An answer in the affirmative might be closer to the truth if the question were worded: Have the limb tissues of the non-regenerating vertebrates evolved in such a way that the work which nerves do has become more difficult? Changes which occur in anuran tissues at metamorphosis and which cause failure to regenerate are reviewed in an excellent article by Polejaiev in the *Biological Reviews* for 1946.[28] It might be possible to overcome these changes in some of the non-regenerating forms by increasing the quantity of nerve fibers in the limb. There has already been initial success along this line. Singer reported in a discussion, late in 1946, at the A.A.A.S. meetings in Boston that he had stimulated partial limb regeneration in adult frogs by deviating the sciatic nerve to the forelimb. We are awaiting the results of our own attempts to stimulate limb regeneration in lizards by increasing the amount of neural tissue.

Returning again to the salamanders, we find that the need for nerves is not over as soon as a blastema has formed, as demonstrated by Weiss,[67] Schotté,[44, 46, 47] and Samarajew.[38, 39] If nerves are sectioned in a salamander limb during the early stages of blastema formation, the blastema regresses. Even after a regenerate has advanced to the stage when the first form changes are apparent, interference with the nerve supply may prevent the formation of distal structures or cause them to be abnormal.[44, 38, 39] The structures affected are the last to take form. Their formation is ordinarily accompanied by rapid growth. When the nerves


9


10


11

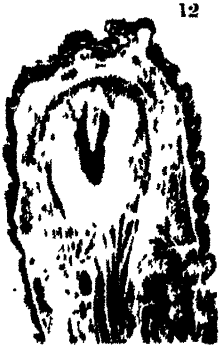

12

FIGURE 9. An untreated *Triturus* limb 23 days after denervation and amputation. In this denervated limb, there had been slight erosion of bony matrix, but this was almost immediately followed by differentiation of cartilage, which may be seen at the extreme right of the photograph along the shaft of the bone. There is another small island of differentiating cartilage to the left and near the middle of the bone on the left. X70.

FIGURE 10. An untreated *Triturus* limb 4 months after a reamputation which had been performed 23 days after the original amputation and denervation. This limb, which probably had some regenerated nerve fibers at the time of reamputation, has formed a cartilaginous callus covered by fibrous scar tissue. X35.

FIGURE 11. A normal untreated stump of a frog's forearm. Here, as in FIGURE 10, can be seen the cartilaginous callus and the fibrous scar tissue. X22.

FIGURE 12. A normal untreated stump of a lizard forearm. This naturally non-regenerating limb shows the same arrangement of cartilaginous callus and scar tissue as seen in the non-regenerating frog's limb (FIGURE 11) and in the non-regenerating partially innervated newt limb (FIGURE 10). X12.

are sectioned, the growth rate decreases markedly. It seems likely that these distal abnormalities result from failure of growth at this critical period. Differentiation unaccompanied by rapid growth also seems to be the cause of distal abnormalities in induced anuran regenerates.

The various treatments so far used to artificially stimulate regeneration have provided large amounts of dedifferentiated material but have not increased growth during the period of finger formation when growth is so necessary a part of morphogenesis. Again, we see the adult anuran limb failing in a process which is supported in the urodeles by nerves.

As mentioned earlier, the conditions of failure to regenerate are different in young larval limbs after denervation. Contact between wound epithelium and internal tissues is made, but no fibrous membrane develops under the wound epithelium. Instead, the wound epithelium continues to exert its histolytic action until the entire limb has disappeared. As in adult limbs, there is no growth of the products of dedifferentiation into a regenerate in the absence of nerves. However, a regenerate forms from these dedifferentiated tissues very rapidly when nerves re-enter the region.[47] A partial explanation of the difference in response of young and older animals is that older tissues have acquired a greater propensity for fiber formation.

These differences between young and old denervated limbs are paralleled by those exhibited in x-rayed limbs. Sufficient dosage to prevent regeneration of larval limbs is followed by dedifferentiation and resorption of entire limbs.[3] Not so with older limbs; x-rayed and amputated limbs of large axolotls soon develop a scar beneath the epithelium and fail to regenerate. The argument that a fibrous scar acts as a barrier between epithelium and internal tissues, thereby preventing dedifferentiation, is supported by two observations from the irradiation experiments. X-rayed appendages of axolotls which had been amputated and had already formed a scar beneath the wound epithelium, were made to regenerate normally by removing the scar.[18] The second observation is that young limbs x-rayed twelve days after amputation do develop a fibrous layer beneath the wound epithelium and do not dedifferentiate.[4] In this case, a large blastema was already present and its cells were about to differentiate when the limb was x-rayed.

No theory of nerve action during regeneration can stand unless it can encompass two reports which are, at first, disconcerting, namely, that after transplantation of limbs nerves are not necessary for regeneration. It would seem unwise to doubt these reports, coming as they did, independently, from two of the best laboratories for the study of regeneration. If we accept them as true, and I think we must, our theory must fall unless transplantation itself causes the changes in a limb which we attribute to nerve action. It is fortunate that Polejaiev transplanted tadpole limbs of the same stage as Schotté and Harland denervated. Polejaiev reported regeneration in the transplanted limbs before nerves had an opportunity to re-enter them,[25] while Schotté and Harland, without transplanting, found no regeneration after denervation.[48]

What does transplantation do to a limb? The answer is found in a quotation from Polejaiev and Ginzburg[30] (see also[17]): "In a control at stage IIa subjected to simple limb amputation, the tissue differentiation is preserved in its essential features: the skeleton does not disintegrate, the muscle patterns maintain their integrity; mesodermal cells of the regeneration rudiment do not accumulate on the amputation wound surface under the epithelium. In case of limb autotransplantation, dedifferentiation of tissues takes place: cartilaginous epiphyses are transformed into dense mesenchymatous thickenings; muscle patterns disintegrate, forming a homogeneous mass of mesenchymatous cells which, assuming a round shape and moving towards the epithelium of the amputation wound surface, form the mesodermal part of the regenerating rudiment."

This is strong supporting evidence for the first part of the theory, namely, that nerves act by causing dedifferentiation (except in very young tissues).

Schneider has also reported regeneration without nerves after transplantation.[40] By transplanting a piece of axolotl limb to the region of the back, regeneration of poor feet was obtained. He also stated that regeneration was faster and better when the sciatic nerve was deviated along with the transplant. These facts are in no way incompatible with the second part of the theory, namely, that nerves also support regeneration by supporting growth without premature differentiation.

Literature Cited

1. Barber, L. W. 1944. Correlations between wound healing and regeneration in fore-limbs and tails of lizards. Anat. Rec. 89: 441–453.
2. Bovet, D. 1930. Les territoires de régénération; leurs propriétés étudiées par la méthode de déviation du nerf. Rev. Suisse Zool. 37: 83–146.
3. Butler, E. G. 1933. The effects of x-radiation on the regeneration of the fore limb of *Amblystoma* larvae. J. Exp. Zool. 65: 271–313.
4. Butler, E. G., & W. O. Puckett. 1940. Studies on cellular interaction during limb regeneration in *Amblystoma*. J. Exp. Zool. 84: 223–239.
5. Butler, E. G., & O. E. Schotté. 1941. Histological alterations in denervated non-regenerating limbs of urodele larvae. J. Exp. Zool. 88: 307–341.
6. David, L. 1934. La contribution du matériel cartilagineux et osseux au blastème de régénération des membres chez les Amphibiens Urodèles. Arch. Anat. microsc. 30: 217–234.
7. Gidge, N. M., & S. M. Rose. 1944. The role of larval skin in promoting limb regeneration in adult Anura. J. Exp. Zool. 97: 71–93.
8. Godlewski, E. 1928. Untersuchungen über Auslösung und Hemmung der Regeneration beim Axolotl. Arch. Entw.-mech. 114: 108–143.
9. Goldfarb, A. J. 1909. The influence of the nervous system in regeneration. J. Exp. Zool. 7: 643–722.
10. Goodwin, P. A. 1946. A comparison of regeneration rates and metamorphosis in *Triturus* and *Amblystoma*. Growth 10: 75–87.
11. Guyénot, E., & O. Schotté. 1926. Demonstration de l'existence de territoires spécifiques de régénération par la méthode de la déviation des troncs nerveux. C. R. Soc. Biol. 94: 1050–1052.
12. Hamburger, V. 1928. Die Entwicklung experimentell erzeugter nervenloser und schwach innervierter Extremitäten von Anuren. Arch. Entw.-mech. 114: 272–363.

13. Harrison, R. G. 1904. An experimental study of the relation of the nervous system to the developing musculature in the embryo of the frog. Am. J. Anat. 3: 197–220.

13a. Hellmich, W. G. 1930. Untersuchungen über Herkunft und Determination des regenerativen Materials bei Amphibien. Arch. Entw.-mech. 121: 135–203.

14. Hines, C. W. 1905. The influence of the nerve on regeneration of the leg of Diemyctilus. Biol. Bull. 10: 44–47.

15. Jeffimoff, M. I. 1931. Die Materialien zur Erlernung der Gesetzmässigkeit in den Erscheinungen der Regeneration. Z. Exp. Biol. (russ.) 7: 352–367 (reviewed by Polejaiev & Faworina, 1935).

16. Jeffimoff, M. I. 1933. Die Rolle der Haut im Prozess der Regeneration eines Organs beim Axolotl. Z. Biol. (russ.) 2 (reviewed by Polejaiev & Faworina, 1935).

17. Liosner, L. D. 1931. Über den Mechanismus des Verlusts der Regenerationsfähigkeit während der Entwicklung der Kaulquappen von *Rana temporaria*. Arch. Entw.-mech. 124: 571–583.

18. Litschko, E. J. 1934. Einwirkung der Röntgenstrahlen auf die Regeneration der Extremitäten des Schwanzes und der Dorsalflosse beim Axolotl. Trudy Lab. Eks. Zool. I. Morf. Jiv. 3: 136–139 (German summary).

19. Locatelli, P. 1925. Formation de membres surnuméraires. C. R. Assoc. Anat. 20: 279–282.

20. Locatelli, P. 1929. Der Einfluss des Nervensystems auf die Regeneration. Arch. Entw.-mech. 114: 686–770.

21. Needham, J. 1942. *Biochemistry and Morphogenesis.* Cambridge Univ. Press. Cambridge, England.

21a. Neukomm, S. 1941. Le centre organisateur dans la régénération des Amphibiens. Rev. Suisse Zool. 48: 519–522.

22. Oliver, M. P., & S. M. Rose. 1946. Induced regenerative processes in denervated limbs of *Triturus*. Anat. Rec. 96: 28.

23. Polejaiev, L. W. 1936. Die Rolle des Epithels bei der Regeneration und in der normalen Ontogenese der Extremitäten bei Amphibien. Zool. Zhurnal 15: 291.

24. Polejaiev, L. W. 1939a. Über die Bedeutung des Epithels und Mesoderms beim Verlust der Regenerationsfähigkeit der Extremitäten bei den Anuren. C. R. Dok. Acad. Sci. URSS 25: 538–542.

25. Polejaiev, L. W. 1939b. Über die Bedeutung des Nervensystems bei der Regeneration der Extremitäten bei den Anuren. C. R. Dok. Acad. Sci. URSS 25: 543–546.

26. Polejaiev, L. W. 1945a. Chemical methods for restoring the regenerative capacity of limbs in tadpoles. C. R. Dok. Acad. Sci. URSS 48: 216–220.

27. Polejaiev, L. W. 1945b. Limb regeneration in adult frog. C. R. Dok. Acad. Sci. URSS 49: 609–612.

28. Polejaiev, L. W. 1946. The loss and restoration of regenerative capacity in the limbs of tailless amphibia. Biol. Rev. Cambridge 21: 141–147.

29. Polejaiev, L. W., & W. N. Faworina. 1939. Über die Rolle des Epithels in den anfänglichen Entwicklungsstadien einer Regenerationsanlage der Extremität beim Axolotl. Arch. Entw.-mech. 133: 701–727.

30. Polejaiev, L. W., & G. I. Ginzburg. 1939. Studies by the method of transplantation on the loss and restoration of the regenerative power in the tailless amphibian limbs. C. R. Dok. Acad. Sci. URSS 23: 733–737.

31. Polejaiev, L. W., & G. I. Ginzburg. 1943. Investigation of ways of formation of regeneration blastema based on calculation of mitotic coefficient. C.R.Dok. Acad. Sci. URSS 43: 315–317.

32. Rose, S. M. 1942. A method for inducing limb regeneration in adult Anura. Proc. Soc. Exp. Biol. & Med. 49: 408–410.

33. Rose, S. M. 1944a. Methods of initiating limb regeneration in adult Anura. J. Exp. Zool. 95: 149–170.

34. Rose, S. M. 1944b. Causes for loss of regenerative power in adult Anura. Anat. Rec. 89: 6.

35. Rose, S. M. 1945. The effect of NaCl in stimulating regeneration of limbs of frogs. J. Morphol. 77: 119–139.

36. Rose, S. M., & C. J. Seller. 1946. Type of regeneration in limbs of frogs after transplantation of adult skin. Anat. Rec. 94: 73.

37. Rubin, R. 1903. Versuche über die Beziehung des Nervensystems zur Regeneration bei Amphibien. Arch. Entw.-mech. 16: 21–75.

38. Samarajew, V. N. 1939a. Denervation of extremity at various stages of regeneration. I. Growth and differentiation of denervated regenerates. Bull. Biol. Med. Exp. URSS 8: 505–508.

39. Samarajew, V. N. 1939b. Denervation of extremity at various stages of regeneration. II. Studies of denervated and control regenerates. Bull. Biol. Med. Exp. 8: 509–512.

40. Schneider, G. 1940. Der Einfluss des Nervensystems auf die Regeneration der Gliedmassen der Axolotl. Bull. Acad. Sci. URSS (Série Biol., German summary, p. 403).

41. Schotté, O. 1922a. La régénération est-elle liée à l'innervation motrice ou à l'innervation sensible? C. R. Soc. Phys. & Hist. nat. Genève 39: 134–137.

42. Schotté, O. 1922b. Le Grand Sympathique—élément essentiel de l'influence du système nerveux sur la régénération des pattes de Tritons. C. R. Soc. Phys. & Hist. nat. Genève 39: 137–140.

43. Schotté, O. 1924. Le Grand Sympathique est le seul facteur nerveux dans la régénération des membres de Tritons. C. R. Soc. Phys. & Hist. nat. Genève 41: 45–52.

44. Schotté, O. 1926. Système nerveux et régénération chez le Triton. Rev. Suisse Zool. 33: 1 211.

45. Schotté, O. 1940. The origin and morphogenetic potencies of regenerates. Growth (Suppl.): 59–76.

46. Schotté, O. E., & F. G. Butler 1941. Morphological effects of denervation and amputation of limbs in urodele larvae. J. Exp. Zool. 87: 279–322.

47. Schotté, O. E., & F. G. Butler. 1944. Phases in regeneration of the urodele limb and their dependence on the nervous system. J. Exp. Zool. 97: 95–121.

48. Schotté, O. E., & M. Harland. 1943a. Effects of denervation and amputation of hindlimbs in Anuran tadpoles. J. Exp. Zool. 93: 453–493.

49. Schotté, O. E., & M. Harland. 1943b. Amputation level and regeneration in limbs of late *Rana clamitans* tadpoles. J. Morphol. 73: 329-363.

50. Schotté, O. E., & M. Harland. 1943c. Effects of blastema transplantations on regeneration processes of limbs in *Amblystoma* larvae. Anat. Rec. 87: 165–180.

51. Schotté, O. E., & A. G. Karczmar. 1944. Limb parameters and regression rates in denervated amputated limbs of urodele larvae. J. Exp. Zool. 97: 43–70.

52. Singer, M. 1942a. The sympathetics of the brachial region of the urodele, *Triturus*. J. Comp. Neurol. 76: 119–143.

53. Singer, M. 1942b. The nervous system and regeneration of the forelimb of adult *Triturus*. I. The role of the sympathetics. J. Exp. Zool. 90: 377–399.

54. Singer, M. 1943. II. The role of the sensory supply. J. Exp. Zool. 92: 297–315.

55. Singer, M. 1945. III. The role of the motor supply, including a note on the anatomy of the brachial spinal nerve roots. J. Exp. Zool. 98: 1–21.

56. Singer, M. 1946a. IV. The stimulating action of a regenerated motor supply. J. Exp. Zool. 101: 221–240.

57. Singer, M. 1946b. V. The influence of number of nerve fibers, including a quantitative study of limb innervation. J. Exp. Zool. 101: 299–338.

58. Taube, E. 1921. Regeneration mit Beteiligung ortsfremder Haut bei Tritonen. Arch. Entw.-mech. 49: 269–315.

59. Thornton, C. S. 1942. Studies on the origin of the regeneration blastema in *Triturus viridescens*. J. Exp. Zool. 89: 375–390.

60. Todd, T. J. 1823. On the process of reproduction of the members of the aquatic salamander. Quart. J. Lit., Sci. & Arts 31: 84–96.

61. Tornier, G. 1906. Kampf der Gewebe im Regenerat bei Begünstigung der Hautregeneration. Arch. Entw.-mech. 22: 348-369.

62. Umanski, E. 1937. Untersuchung des Regenerationsvorganges bei Amphibien mittels Ausschaltung der einzelnen Gewebe durch Rontgenbestrahlung. Biol. Zhurn. URSS 6: 757-758 (German summary).

63. Umanski, E. 1938. The regeneration potencies of axolotl skin studied by means of exclusion of the regeneration capacity of tissues through exposure to x-rays. Bull. Biol. Med. Exp. U.S.S.R. 6: 141-145.

64. Walter, F. K. 1912. Welche Bedeutung hat das Nervensystem fur die Regeneration der Tritonextremitaten? Arch. Entw.-mech. 33: 274-296.

65. Wolff, G. 1902. Die physiologische Grundlage der Lehre von den Degenerationszeichen. Virch. Arch. 169: 308-331.

66. Wolff, G. 1910. Regeneration und Nervensystem. Festschrift Richard Hertwig. 3: 67-80.

67. Weiss, P. 1925. Abhängigkeit der Regeneration entwickelter Amphibienextremitäten vom Nervensystem. Arch. Entw.-mech. 104: 317-358.

Discussion of the Paper

Dr. E. G. Butler (*Princeton University, Princeton, N. J.*):

One of the points I should like to discuss is whether, at the present time, we have final convincing evidence that regression does not occur in amputated nerveless limbs of adult urodeles. It was my understanding from listening to Dr. Rose that he has found no regression in the nerveless limbs in his experiments. On the other hand, I believe that in unpublished experiments Dr. Schotté has found that regression occurs in adult denervated limbs after amputation and prolonged absence of nerves. I should like to ask Dr. Rose how long he maintained the limbs in his experiments in a nerveless condition. It seems to me that, in this problem, time is a very important factor. It is not so much a question whether regression in an amputated limb takes place during a 30- or 60-day period of nervelessness, as whether it will take place during a 4- or 6-month period. Certainly, extreme regression takes place in larval limbs. Can it be that the situation is so different between the larva and the adult? In my opinion, we should have data from experiments on adults carried over a long period of time before we endeavor to answer this question positively.

It is my understanding from Dr. Rose's paper that he regards the epidermis as contributing to the blastema. It is possible that cells from the epidermis pass into the blastema, but do they form a functional component of the blastema? I ask this question not as a criticism of Dr. Rose's work, but rather for information. One of the most involved problems with which I am acquainted is that of endeavoring to recognize the source of cells which contribute to the regeneration blastema. Is there clear evidence that cells from the epidermis not only enter the blastema region, but also actually take part in regenerative activity?

Lastly, I should like to mention the matter of blastema age and its relation to limb regression. My work on regeneration has been primarily with urodele larvae and I am basing my statements chiefly on these.

I feel that the time has come when we can no longer simply refer to a regeneration "blastema." A newly formed blastema is a far different structure, morphologically and physiologically, than an older one. The blastema is a continually changing structure and at different times exhibits different types of activity. So far as regression is concerned, a young blastema is incapable of preventing a nerveless larval limb from undergoing regression; the presence of an older blastema, however, prevents a limb from regressing. I should like to suggest, therefore, that, in studying the physiology of regeneration, in adults as well as in larvae, the age of the blastema always be taken into account. It seems to me that this is an important consideration in the type of blastema which Dr. Rose is studying in his salted limbs.

DR. S. M. ROSE:

We have kept animals whose limbs had been denervated and amputated for as long as five months. Some of these animals began to regenerate limbs after several months' delay, others formed the type of callus shown in FIGURE 10, and still others formed practically no new tissue. There was no measurable regression in any of these animals except in one case, and in that only after the subclavian artery had been inadvertently severed. Judging with the work of Walter and of Singer as a basis, I think that these groups are different because of difference in nerve number. They are alike in failing to regress. Some other factor seems to be more important in causing regression than a nerve number below that necessary for limb regeneration. It would be very interesting to know whether adult limbs which had been maintained absolutely nerveless for a long time would, in time, begin to regress. As I remember, Dr. Schotté did believe that nervelessness for a long time would be followed by regression. Recently, in a discussion, he stated that regression had been obtained in adults after denervation if the amputation was performed through a regenerate. He did not speak of regression of older denervated tissues. I think that the question of the amount of regression after long periods of nervelessness is still unsettled and that we must await a projected paper by Dr. Schotté which will more fully cover the problem of delayed regression.

In our present work, it is clear that nerves play an important role in dedifferentiation. Without them, there is scarcely any dedifferentiation. This is very different from the situation in young larvae. I do believe that there are real differences between larva and adult; both the denervation and x-ray radiation experiments indicate it. My guess is that a difference originally quantitative, which we see reflected in the ability to produce collagen, is the basis for the resulting qualitative difference between young larva and adult. An initial study of some of the general age changes as they affect regeneration has been made by Miss Goodwin.[10] Much more work on this important problem is needed.

I do not know whether epidermal cells enter the blastema and later differentiate into "mesodermal" tissues. There is direct evidence that epidermal cells pass into the blastema. This has been described by Godlewski,[8] by Hellmich,[13a] by Neukomm,[21a] and by Rose.[35] Godlewski, Neukomm and I were almost sure that we could see intergradations between epidermal cells and cells of the blastema. There is also the experiment by Umanski[63] indicating that a blastema may form from untreated skin transplanted to an x-rayed limb. That skin contained both epidermis and dermis, either or both of which can be suspected. We have unpublished studies from which it appears certain that vitally stained cells leave the epidermis and enter the blastema. Later, a few chondrocytes can be found containing the vital dye. Again, uncertainty creeps in and this cannot be considered evidence that epidermal cells become chondrocytes, because many of the epidermal cells are phagocytized and their stained granules transferred to their mesodermal predators. Hellmich, without using vital dyes, came to the conclusion that all of the epidermal cells in a blastema are destroyed. That may be true, but the observations above and other considerations[35] have kept alive the unproven idea that the epidermis may be one of the sources of blastema cells. A better test of the idea is needed than any used so far.

I fully agree with Dr. Butler that consideration of the age of a blastema is important. I should also like to say that knowledge of biochemical changes during this period would be very useful.

QUANTITATIVE STUDIES ON LOCOMOTOR RESPONSES IN *AMBLYSTOMA* LARVAE FOLLOWING SURGICAL ALTERATIONS IN THE NERVOUS SYSTEM

By S. R. DETWILER
Department of Anatomy, College of Physicians and Surgeons,
Columbia University, New York, N. Y.

Introduction

OVER an extended period of years, the author has been interested in the capacity of various portions of the embryonic nervous system of *Amblystoma* to undergo structural and functional readjustments when subjected to new conditions by means of grafting.

One of the earliest experiments in this connection consisted in substituting trunk spinal segments (6, 7, and 8) from a donor embryo for the brachial segments (3, 4, and 5) of a host embryo. The embryos ranged in age from those with completely closed neural folds (stage 21) to those with a prominent tail bud (stage 30). A histological study of the cords in host larvae, approximately 50 days after the operation, showed that the grafted segments in the new position had undergone a volume increase and cellular hyperplasia which approximated the size and cellular content normally characterizing the brachial region of the cord. The forelimbs in 50 per cent of the cases exhibited normal function and were supplied by a typical brachial plexus, whose segmental contribution took origin from the grafted segments. These and other experiments dealing with the interchange of various spinal segments and the effects upon cellular proliferation, showed that the number of cells normally characterizing a given region is not inherently and irrevocably fixed. Some regions, however, are more plastic than others. Some have a greater "potential" for proliferation than others, regardless of their axial position.*

In some recent experiments (Detwiler, 1945) upon the embryonic brain, the presumptive hemispheres were excised in order to study any possible morphogenetic effects of their absence upon the development of the remainder of the brain, particularly the medulla. It had been shown by Burr (1916a) that the cerebral hemispheres of *Amblystoma* are dependent upon the intactness of the nasal sacs for their complete development. He also showed (1916b) that, when a cerebral hemisphere is removed from embryos,* it will regenerate only if the nasal placode is left intact. The stimulus for regeneration was assigned to the ingrowth of the olfactory nerve fibers, a contention which received support by later experiments (1930).

---

* For a discussion of the morphogenetic responses of various spinal segments in altered positions see Detwiler, 1936. Hamburger (1946) has studied the proliferation of cells in mechanically isolated portions of the chick spinal cord, and has discussed the results in relation to earlier findings on the amphibian cord following certain experimental procedures.

My own experiments upon the forebrain corroborated Burr's results in that there was no evidence of regeneration of the hemispheres. In these experiments, which were conducted on Harrison's stage 21±,† the excised region included the presumptive nasal placodes and the eye rudiments. Both unilateral and bilateral excisions were made. Since it was doubtful whether embryos lacking the entire forebrain, eyes, and nasal placodes could lead an independent existence beyond the yolk resorption stage, many were fused parabiotically with normal embryos to serve as nurses and at the same time as controls for the operated components. Others were allowed to develop as "free" individuals.

External malformations in the shape of the head became visible fairly early. The growth of the upper jaw was greatly reduced, resulting in a marked protrusion of the lower jaw (FIGURES 1 and 2). In cases of unilateral excision, this condition existed only on the side of operation (FIGURE 3).

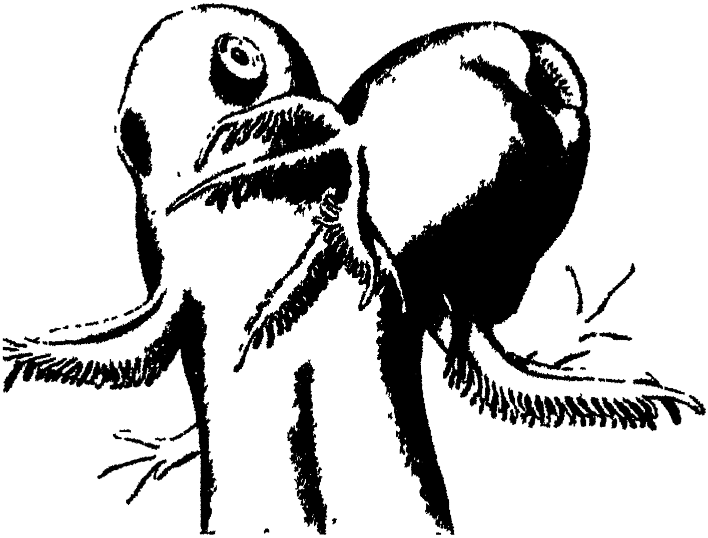


FIGURE 1. Parabiotic twin (FBEBP20) 26 days after operation. ×8. The right component lacks both cerebral hemispheres, eyes, and nasal organs.

When the larvae had reached the feeding stage, most of the nonparabiosed individuals, despite the absence of forebrain, eyes, and nasal placode, exhibited snapping reactions when a needle was moved gently

---

* Stages of operation not indicated. † At stage 21 the neural folds are completely closed.

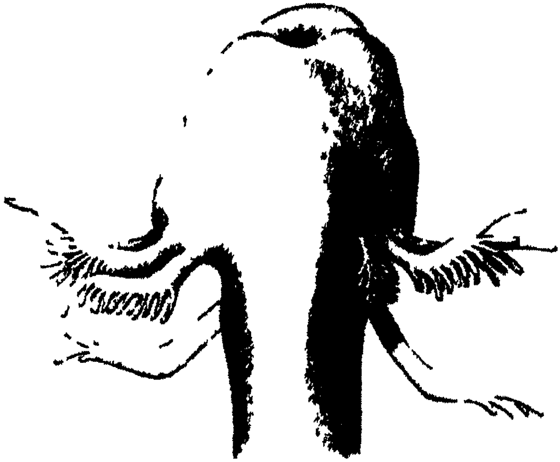


FIGURE 2 *Amblystoma* larva (FBEB44) 26 days after operation ×8 The forebrain, eyes, and nasal placodes were removed from the embryos in stage 21

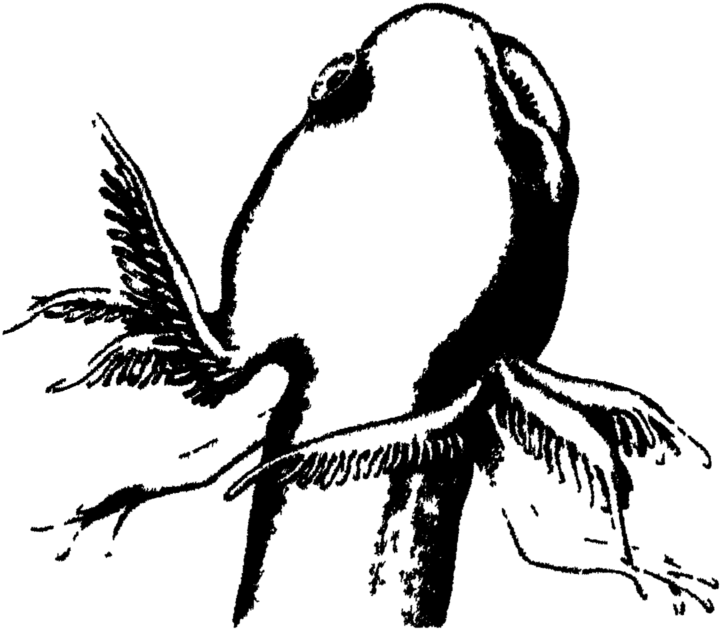
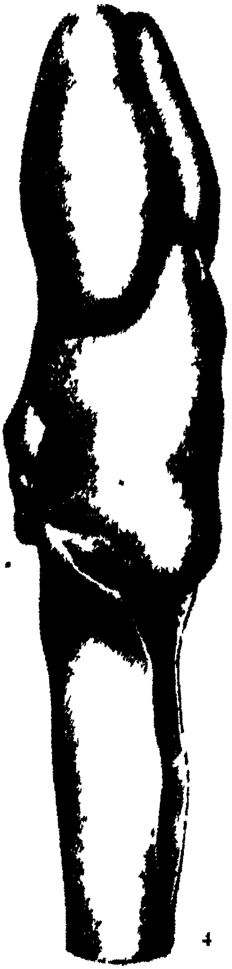
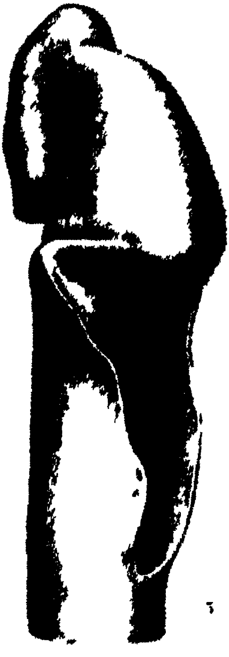


FIGURE 3 *Amblystoma* larva (FBE21) with absence of the right hemisphere, the ipselateral eye, and nasal sac, 30 days after operation × 8

back and forth along the side of the head. They were capable also of feeding upon daphnia and small enchytraeid worms Sharrer (1932) had indicated previously the significance of the lateral line sense organs in the characteristic snapping reactions. It had also been shown later (Detwiler and Copenhaver, 1940) that the growth rate of larvae lacking eyes and nasal organs may keep pace with that of controls under conditions of maximal feeding. This indicated clearly that the lateral line sense organs alone may constitute an adequate receptor apparatus for the detection of food in motion.



4



5

FIGURE 4 Wax reconstruction of the brain (dorso-lateral view) of the normal component of parabiotic twin FBEB22 x20

FIGURE 5 Wax reconstruction of the brain (dorso-lateral view) of the operated component of parabiotic twin FBEB22 x20 Note lack of hemispheres and the dorsal portion of the diencephalon

The intake of food, and the growth of the larvae lacking the forebrain, eyes, and nasal placodes, was markedly curtailed as compared with the controls. Spontaneous behavior in general was greatly reduced, both quantitatively and qualitatively, especially the foraging reactions, regardless of the amount of food available. The mean length of the larvae at 45 days of age was slightly more than 50 per cent of the length of maximally fed normal animals of similar age. Nevertheless, the motor activities concerned with lurching, engulfing food, chewing, and swallowing were carried out in an integrated manner, although they were decidedly less vigorous than in larvae with intact hemispheres. In the parabiotic twins, the eating function was taken over largely by the more active conjoined normal component, but here, too, the operated component was seen to snap and engulf food. Although actual measurements were not recorded for all cases, it may be said that the growth of the twins was greater than that of the operated "free" individuals, but not as great as in the normal control larvae.

Wax reconstructions of the brains of the 2 components in twin 22 are shown in FIGURES 4 and 5. The operated component not only lacks the hemispheres, but the dorsal portion of the diencephalon is also wanting. The hypothalamic region and the pituitary gland (both glandular and neural portions) are present and essentially normal. This situation was characteristic of all cases whether the individual was parabiosed or not. The defective diencephalon is interpreted as resulting from partial ablation during the operation rather than to any morphogenetic influence due to the absence of the hemispheres.

Despite the complete absence of the hemispheres, an incomplete diencephalon, and a possibly smaller mesencephalon, the medulla in all cases showed no significant reduction in size. The medullas of the operated components were somewhat shorter than those of the normal components of the twins (FIGURE 6, *cf.* *A* and *B*), but their volume in all cases studied, except one, ranged from 92 to 99 per cent of normal (Detwiler, 1945, Table 1).

In cases with unilateral excision of the forebrain, there occurred a compensatory enlargement of the contralateral nasal sac and the adjacent hemisphere. The single hemispheres in 4 cases studied showed volume increase ranging from 73 to 88 per cent of the volume of both hemispheres in a control larva of similar length (Detwiler, *op. cit.*, Table 2). A histological study of nuclear size forced the conclusion that the compensatory enlargement of the hemisphere was accomplished by a cellular hyperplasia.

The capacity of the young larvae lacking hemispheres, eyes, and nasal organ to obtain food and to lead an autonomous existence, led to experiments designed to test the behavior of larvae lacking the midbrain.

Nicholas (1930) excised the embryonic mesencephalon and replaced it with a graft consisting of the forelimb and adjacent pronephric rudiments, thus isolating the forebrain from the remainder of the nervous

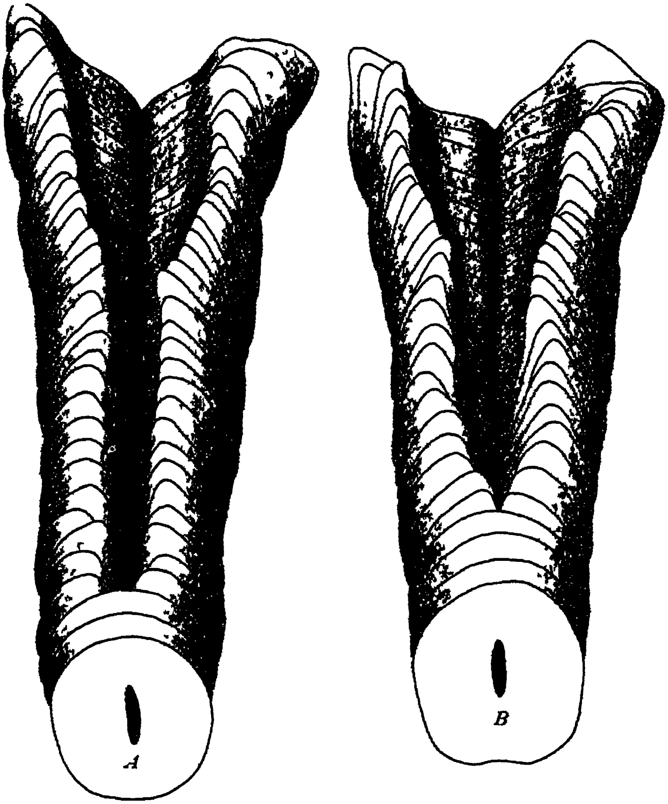


FIGURE 6. Graphic reconstruction of the medullas of normal component (*A*) and operated component (*B*) of parabiotic twin FBEB3. The operated component lacks the cerebral hemispheres, eyes and nasal sacs. ×40. (See FIGURE 1.)

system. He found that the early responses (*C* and *S* reactions) were carried out normally. He said little about the effects of the operation upon behavior of older larvae, except that spontaneous movements were not present and that vigorous swimming movements could be elicited by stimulation. These apparently diminished in amount as the animals became older. Nicholas did observe that his larvae were unable to obtain food, and this failure was regarded by him as due to the fact that "the removal of the mesencephalon created a block in the nervous system so that normal feeding responses could not be obtained."*

Experimental Results

In my own experiments upon the mesencephalon, three types were performed: (A) excision of the mesencephalon from embryos in stage 20–21 and replacement by the first and second spinal cord segments from a donor embryo in stage 22–23; (B) end-to-end reversal of the entire midbrain in embryos of stage 22; and (C) unilateral reversal of the midbrain in embryos in stage 20–21, involving reversal of the anteroposterior axis only.

---

* In cases where the larvae were kept beyond the feeding stage (stage 46), the animals were fed by artificial means.

Excision of the Mesencephalon and Its Replacement by Cord Segments I and II. This procedure differs from that of Nicholas (1930) in that nervous continuity is maintained between the fore- and hindbrain. The majority of the larvae at the feeding stage were unable to eat. Several exhibited feeble snapping reactions but were unable to obtain their prey. Only 2 cases in 24 were recorded as having eaten, and their intake was very meager. In most cases, the jaws were immobile and held slightly apart. This condition readily explains the inability to capture *Daphnia* in those cases where snapping was attempted. The incidence of kyphosis was very low; there were only 2 cases in this series which developed this condition. This is in marked contrast with the results obtained by Nicholas, all of whose larvae developed a marked ventral bending of the head and tail. The fact that the larvae in the present experiments developed as straight individuals naturally favored a critical study of their swimming responses.

Observations upon Swimming Responses. All the early larvae exhibited normal *C* and *S* flexures and swimming responses. In later stages, however, despite previous manifestations of an adequate intraspinal swimming mechanism, the larvae showed progressive failure both in their responses to tactile stimulation and in their locomotor ability. The method of studying behavior was as follows: Individual larvae (normal and mesencephalonectomized of similar stages) were placed in a Syracuse dish and stimulated 25 times at approximately 5-second intervals. This was done by stroking the skin over the myotomes with a human hair, according to the method of Coghill (1909). At the end of 25 stimulations, the number of "misses" was recorded. In addition to studying the responses to tactile stimulation, the distance traveled in response to 25 stimulations was measured. This was done as follows: A circle, the diameter of which equaled the inside diameter of a Syracuse dish, was drawn upon a card. This was divided into 10 sectors of arc (36° each). The sectors were numbered successively from 1 to 10 both clockwise and counterclockwise, and the card was then placed beneath the dish so that the circle coincided with the inner wall of the dish. A larva was placed adjacent to the inner wall of the dish and headed in a clockwise direction. It was then stimulated and the number of sectors traveled before coming to rest recorded. This was done for 25 stimulations at approximately 5-second intervals. At the end of these, the number of positive reactions and the total distance traveled in units (sectors of arc of 36° each) were recorded for each larva. Young larvae are strongly thigmotactic and nearly always swim along the wall of the dish. In order to avoid occasional short-cutting, the device was improved by placing a glass ring within the Syracuse dish which provided a "moat" about 7 mm. in diameter. The improved device is shown in FIGURE 7.

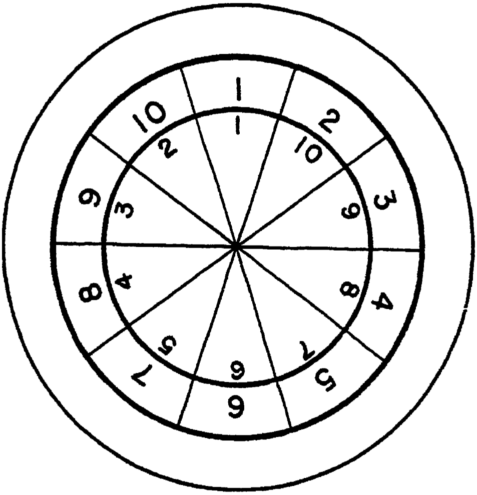


FIGURE 7. An improved device placed beneath a Syracuse dish for quantitating the distance traveled by young *Amblystoma* larvae (stages 39 to 46 and beyond). The outer heavy circle corresponds to the inner wall of the dish; the inner heavy circle indicates a glass ring the height of the dish. The space between the two represents a "moat" approximately 7 mm. in diameter. Each larva was placed in the moat and stimulated 25 successive times at approximately 5-second intervals, and the total distance traveled was recorded in units (sectors of arc). The glass ring has been added to the original device (Detwiler, 1945, Figure 1) to prevent larvae from occasionally short-cutting as they swim along the wall of the dish (v. text).

Records were made upon 2 separate groups each consisting of 4 control and 6 experimental larvae. The data on these 2 groups are shown in FIGURES 8, 9, 10 and 11. FIGURES 8 and 10 show the responses to tactile stimulation, while FIGURES 9 and 11 record the average distance traveled by the larvae at the stages indicated. It is seen that, beyond stage 41, the number of "misses" among the experimental animals was indeed very striking. Although there was some variation in the average distance traveled by the normals at the various stages (solid columns), it is clear that the propulsive power of the experimental larvae was greatly reduced (stippled columns). In no single experimental larva from stage 41 on did the maximum swimming response equal the minimum response of any of the normal larvae.

It is apparent from the findings that the execution of normal locomotor responses quite early in larval life depends upon the intactness of the mesencephalon. Up to about stage 40, the absence of this structure appears to have little or no effect upon the swimming activities, which are autonomous in the sense that they apparently can be carried out entirely independent of any structural connection with the midbrain. The rapid falling-off in the locomotor capacity at this period (stage 40–41) is regarded as evidence of the importance of the midbrain tectum

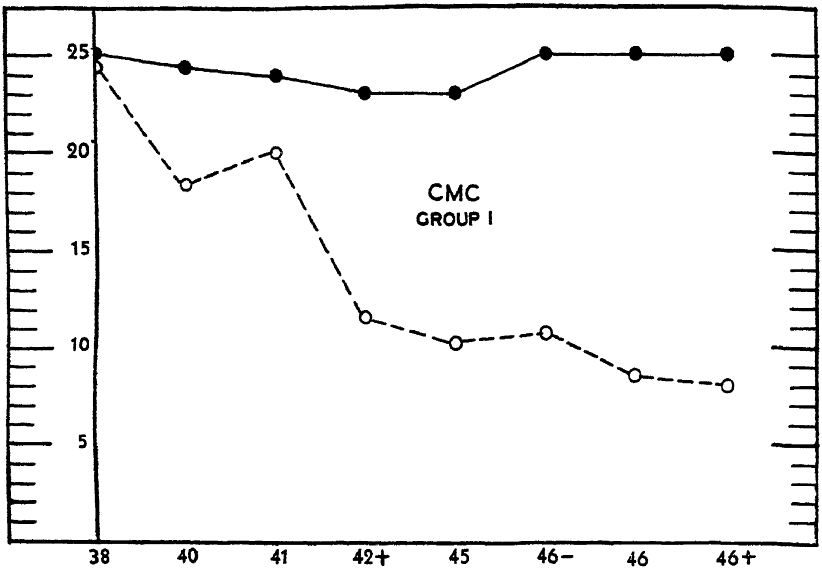


FIGURE 8. Graph showing locomotor responses of *Amblystoma* larvae (CMC-group 1) to 25 successive tactile stimulations at 5-second intervals. Continuous line shows the average for 4 normal control larvae; broken line shows average for 6 larvae in which the midbrain was replaced by the first two segments of the spinal cord. *X* axis represents stage; *Y* axis indicates number of responses.

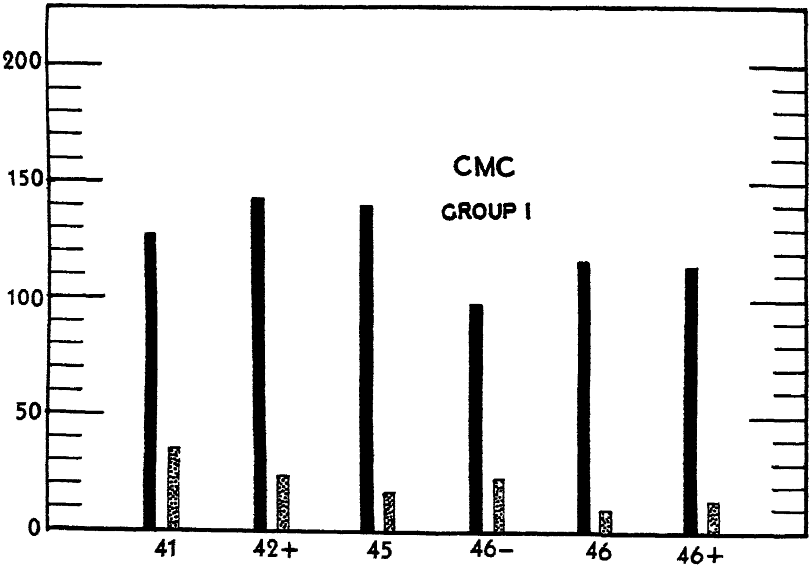


FIGURE 9. Graph showing total distance traveled by the larvae in group 1, in response to 25 successive stimulations at 5-second intervals. The solid columns show the average distance covered by 4 control larvae; the stippled columns show the average distance traveled by 6 larvae whose midbrain was replaced by the first and second spinal cord segments. *X* axis designates stages; *Y* axis represents distance in units (sectors of arc as shown in FIGURE 7).

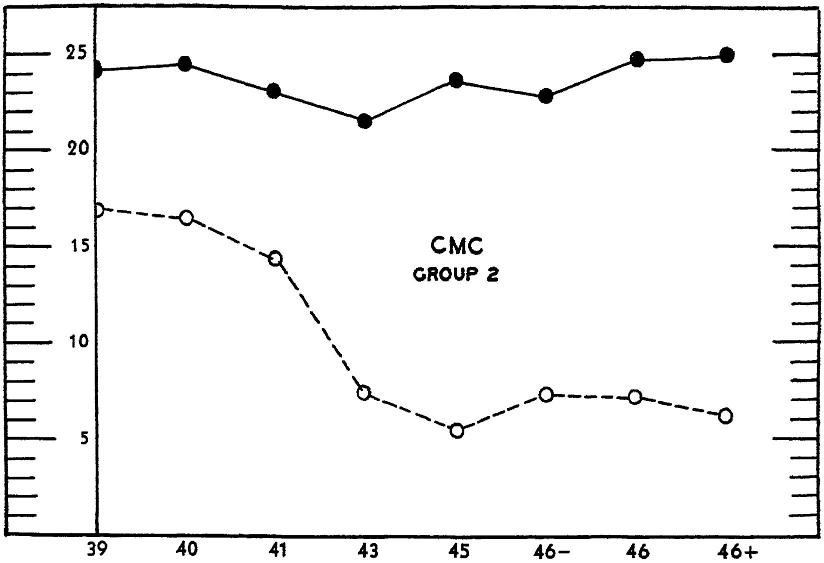


FIGURE 10. Graph showing locomotor responses of the larvae in group 2 (CMC) to 25 successive tactile stimulations at 5-second intervals. Continuous line shows average for 4 control larvae; broken line shows average for 6 larvae with midbrain replaced by spinal cord segments 1 and 2. *X* axis indicates stages; *Y* axis indicates number of responses.

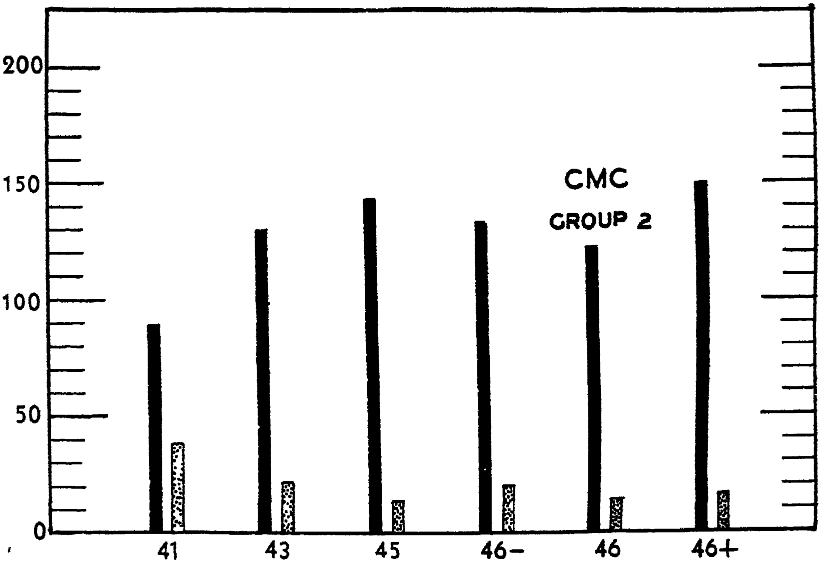


FIGURE 11. Graph showing total distance traveled by the larvae in group 2 (see explanation of FIGURE 9). *X* axis designates stages; *Y* axis represents distance in units (sectors of arc as shown in FIGURE 7).

in coordinating and sustaining these functions. Crossed and uncrossed tracts arising in the tectum invade the medulla as the tecto-bulbar tracts. These make synaptic connections with tracts which arise in the medulla and invade the cord as the bulbo-spinal tracts (Herrick, 1914). However, Herrick (1939) says that some of the fibers arising in the tectum probably invade the spinal cord directly as the tecto-spinal fibers. According to him, the uncrossed tecto-bulbar tracts develop earlier than the crossed fibers of this system. These latter appear in preparations "subsequent to early swimming stages." In the earlier feeding stages, he finds that essentially adult relations have been attained.

It should be pointed out here that larvae lacking the midbrain are still capable of some locomotor activity, *i.e.*, the intra-spinal mechanism can still function, but to a greatly diminished degree (FIGURES 9 and 11). The evidence seems clear, however, that the spinal apparatus, which is entirely adequate for normal propulsion in the early stages (37–40±), later falls under mesencephalic control.

Although no attempt was made to trace out the pre-otic cranial nerves, it would appear that the failure of the animals in this series to eat (*vide supra*) was largely due to the defective development and displacement of the V–VII ganglionic complex as a result of the operation.

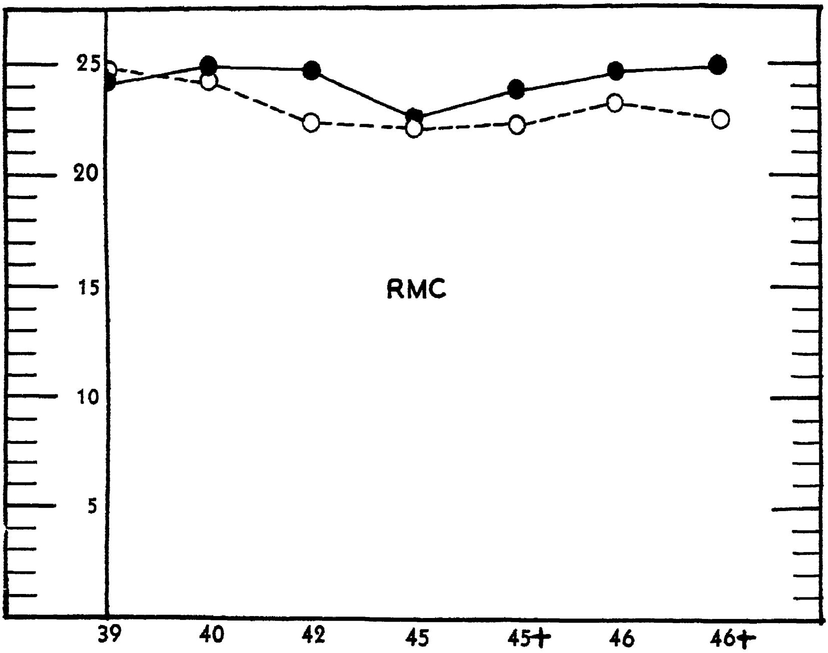


FIGURE 12. Graph showing locomotor responses to 25 successive tactile stimulations at 5-second intervals. Continuous line shows average for 4 normal control larvae; broken line shows average for 6 larvae in which the midbrain was reversed end for end at stage 22. *X* axis designates stages; *Y* axis indicates number of responses.

Failure to cut precisely between embryonic mesencephalon and medulla resulted in the inclusion of the anterior end of the medulla in many excisions—thus involving either the inclusion of or the disturbances to the adjacent V–VII complex.

End-to-End Reversal of the Mesencephalon. As in the series described above, the early *C* and *S* responses were normal. The responses of larvae to tactile stimulation and their motor ability were tested in the manner described above. The results are shown in FIGURES 12 and 13, and they reveal the fact that, despite complete reversal of the midbrain, the larvae responded to tactile stimulation nearly as well as did the controls. More striking, perhaps, is the fact that the motor capacity was only slightly lowered. This is in marked contrast with the greatly curtailed motor capacity of those lacking the midbrain (*cf.* FIGURES 9 and 11). A microscopic examination of the larvae in this series showed that, in 4 of the 9 cases studied, there was no detectable evidence that the midbrain had been reversed (FIGURES 14 and 15). In 5 others, which were incapable of feeding, there were morphological deficiencies which could well account for this defect (Detwiler, 1946a. p. 129).

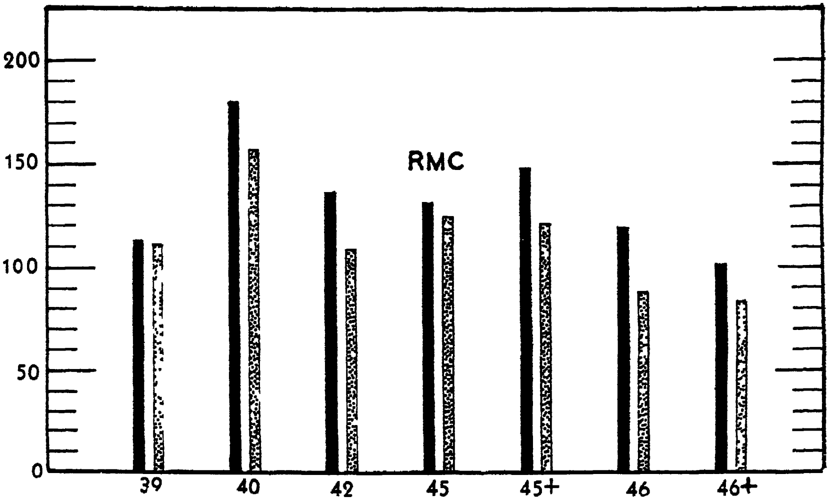
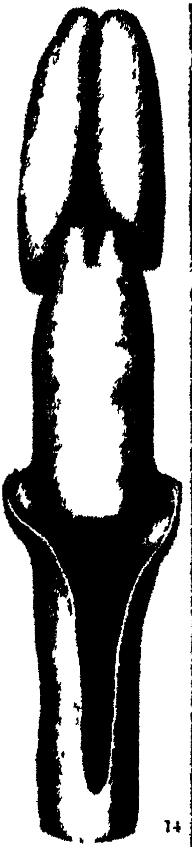


FIGURE 13. Graph showing total distance traveled by larvae in response to 25 successive stimulations at 5-second intervals. The solid columns show the average scores for 4 control larvae; the stippled columns show the average scores of 6 larvae in which the midbrain was reversed end for end in stage 22. *X* axis designates stages; *Y* axis indicates distance in units (sectors of arc as shown in FIGURE 7).

Unilateral Reversal of the Midbrain. In these experiments, the right half of the mesencephalon was excised and replaced by the left half from a donor embryo, thus reversing only the antero-posterior axis. The young larvae, like those of the two previous series, showed normal developmental behavior. No quantitative records were made of the motor activities until stage 46, when 8 larvae were tested along with 6 normal



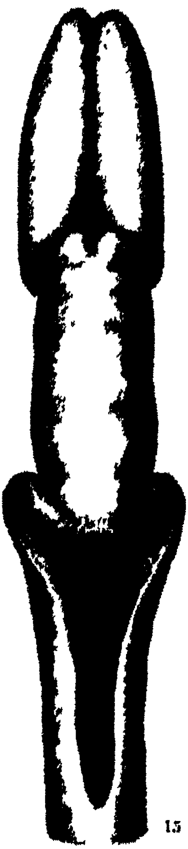


FIGURE 14. Reconstruction model of the brain of a normal larva (48 days, 35 mm.) for comparison with the brain of case RMC60 (with reversed mesencephalon) as shown in FIGURE 15. ×40.

FIGURE 15. Reconstruction model of the brain of case RMC60 (42 days, 38 mm.). The embryonic midbrain was reversed end for end in stage 22 (*cf.* FIGURE 14). ×40.

controls of the same developmental stage. The results showed that the average distance covered by the experimental animals equaled that of the controls. Certain developmental abnormalities were encountered in some of the cases in this series; these have been described previously (Detwiler, 1945, p. 132). One of the most interesting conditions consisted in the failure of the graft to fuse with the cephalic and caudal stump of the host brain. In consequence, the graft developed into an isolated, bilaterally symmetrical whole, and the excised midbrain half was restored by regeneration from the contralateral intact half (Detwiler, 1946b). These cases are of especial interest in showing that, at this stage of development, half a mesencephalon is capable of restoring the whole. The same capacity had been shown previously for the medulla (Detwiler, 1944).

Locomotor Responses of Larvae Lacking the Cerebral Hemispheres. Although the general behavior of larvae lacking the hemispheres had been observed prior to those without the midbrain, no quantitative studies on locomotor responses were made at the time. Consequently, the experiments were repeated and the motor ability tested with the device shown in FIGURE 7. Two groups each consisting of 5 operated and 5 con-

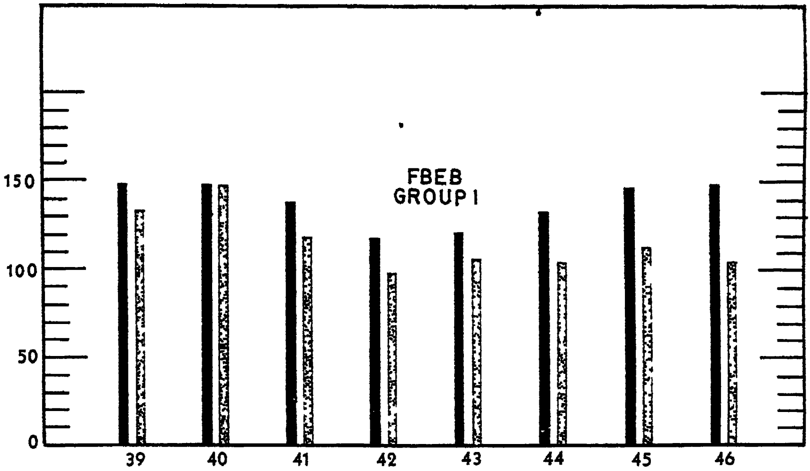


FIGURE 16. Graph showing total distance traveled by the larvae in group 1, in response to 25 successive stimulations at 5-second intervals. The solid columns show the average scores for 5 control larvae; the stippled columns give the average scores for 5 larvae lacking the cerebral hemispheres and the dorsal portion of the diencephalon. *X* axis designates stages; *Y* axis represents distance in units (sectors of arc as shown in FIGURE 7)

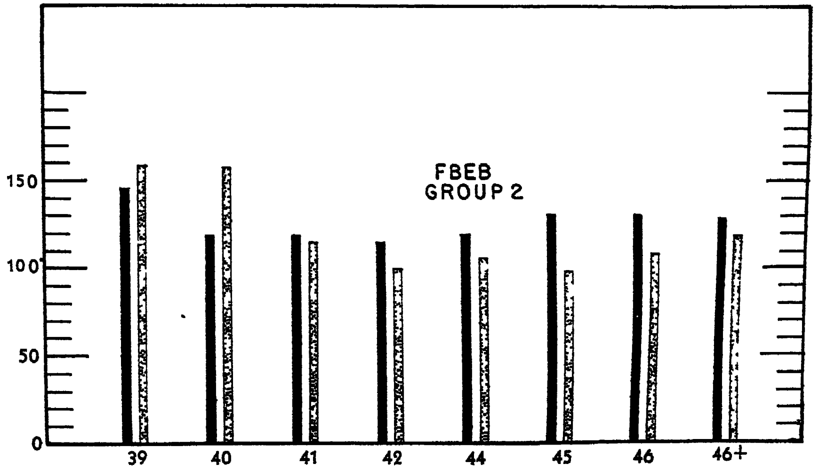
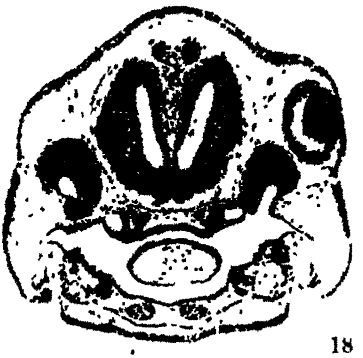
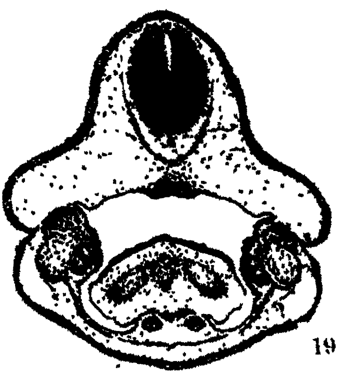


FIGURE 17. Graph showing average scores of 5 control larvae (solid columns) and 5 experimental larvae (stippled columns) in group 2. See explanation of FIGURE 16.

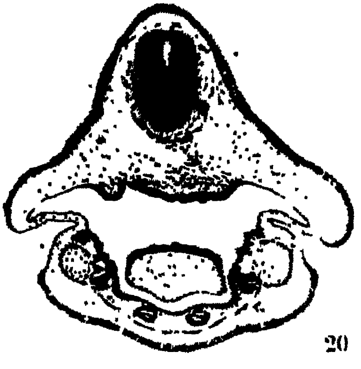
rols were tested from stages 39 to 46+ (feeding stage). The averages of e scores for both normal and operated larvae in the two separate groups are illustrated graphically in FIGURES 16 and 17. These show that the operated larvae in the earlier stages do somewhat better, when compared with their controls, than they do in later stages (stages 45 and 46). The lowered motor ability, however, is only slight as compared with the extreme motor incapacity of larvae lacking the midbrain. The lack of the forebrain also had no effect upon the responses to tactile stimulation. The number of "misses" to trunk tactile stimulation among the operated larvae was no greater than among the controls. This is in marked contrast with the failure exhibited by the mesencephalonecto-



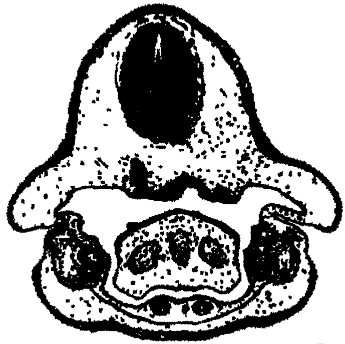
18



19



20



21

FIGURE 18. Photomicrograph of transverse section (2–1–4) of a normal larva (CA) at stage 46 (feeding stage) showing cerebral hemispheres, nasal placodes, and an eye. It is the thirty-first section caudal to the first appearance of the brain in section. x33.3.

FIGURE 19. Photomicrograph of section (1–4–1) from case FBEB (4A) at stage 46. It is the seventh section caudal to the first appearance of the brain in section. The larva is completely devoid of cerebral hemispheres and the dorsal portion of the diencephalon. x33.3.

FIGURE 20. Photomicrograph of section (1–3–4) from case FBEB (25A). It is the seventh section caudal to the first appearance of the brain in section. The larva lacked the hemispheres and the dorsal portion of the diencephalon. x33.3.

IGURE 21. Photomicrograph of section (1–4–2) from case FBEB (25). Conditions same as those deibed for FIGURES 19 and 20. x33.

mized larvae, which showed a striking decline in response to tactile stimulation subsequent to stages 40 and 41.

Morphological Observations. Ten operated larvae and 1 control subsequent to stage 46 were fixed and studied in serial transverse sections. All 10 cases showed a total lack of the hemispheres. In addition, the dorsal portion of the diencephalon was also wanting. The hypothalamic region and pituitary gland were present in most cases, but in several this region was defective.

Four of the sectioned animals are illustrated in FIGURES 18–21. FIGURE 18 is a photomicrograph of the thirty-first section caudal to the beginning of the hemispheres in a control larva (*A*). The hemispheres are prominent, as are also the nasal placodes and an eye. The microphotographs shown in FIGURES 19, 20, and 21 represent the seventh section caudal to the first appearance of brain tissue in the operated animals. The brain in these sections is readily recognizable as mesencephalon.

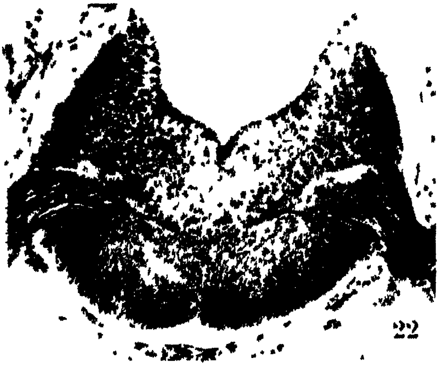
Many of the animals of this series failed to feed. Eight larvae which did eat followed a behavior pattern similar to those described previously (Detwiler, 1945), *viz.*, curtailed spontaneous behavior, rather feeble lurching, reduced food intake, and marked reduction in growth. The activities involved in lurching, engulfing, chewing, and swimming were performed in an integrated manner.

Motor Responses Following the Excision of the Right Mauthner's Neurone and of the Right Ear Vesicle. In teleost fishes and in larval Amphibia, there exist 2 giant neurones known as Mauthner's fibers. These have been of considerable interest to neurohistologists because of certain anatomical features they exhibit—mainly the gigantic size of the cell, the high degree of differentiation, the characteristic position in the medulla, the extensive dendritic connections, the internal cell structure, and the nature of the synapses.

Studies by different authors upon the anatomical connection of these cells in various forms have resulted in several suggestions regarding their function.

In *Amblystoma*, the two perikarya occupy a lateral position in the medulla at the level of the entrance of the VIII nerve (FIGURE 22), and the axones, after decussating in the medulla, course caudally through the spinal cord in a ventral position (FIGURE 23) where connections are made with both motor and intercalary cells (Coghill, 1934).

With regard to the function of this cell, Bartelmez (1915) says: "The reduction of the latent period by elimination of the synapses and the highly medullated character of the system have led me to believe that we are dealing here with a reflex in which speed and precision are very important, and I would suggest that it is this reflex which enables the animal to keep perfect control of its equilibrium in the most rapid and intricate movements." He also suggests that the great caliber of the axones, the broad contact surfaces in the synapse of the lateral dendrit



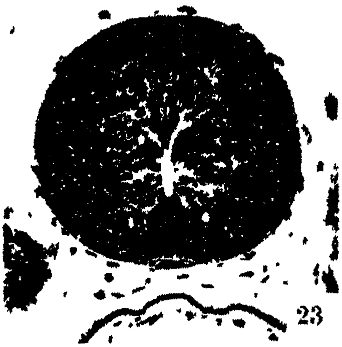


FIGURE 22. Photomicrograph of section (1-3-5) through the medulla of case MPEB (2A) (39 days), showing Mauthner's cells, and their positional relation to the VIII nerve. ×50.

FIGURE 23. Photomicrograph of section through the spinal cord at the level of the third spinal nerve showing position of Mauthner's fibers. ×66.

and the highly insulated character of the whole path, all point to a very rapid reflex. The fact that Mauthner's cell, in addition to its connections with the VIII nerve roots, has connections with every center in the brain which receives impulses that may be used in equilibrium, strongly suggested to Bartelmez an important role in equilibratory reflexes.

Although much attention had been paid to the dendrite connections of the cell body in the brain, it was not until the observations of Coghill (*op. cit.*) that important and hitherto unknown relations of the axonal collaterals in the cord became known. Based on his studies, Mauthner's fibers are regarded as constituting an inhibitory system. Following his observations on the synaptic connections within the cord, Coghill says: "Upon the hypothesis that the action is inhibitory, a volley of impulses passing cephalo-caudal along Mauthner's fibers would inhibit local sensory-motor responses, represented by the cells of intercalated type, in favor of the total action pattern and at the same time would inhibit the antagonist of one side in favor of the agonist of the other, as represented by the cells of motor type. Such inhibition would facilitate the cephalocaudal progressing flexures which effect swimming. Possible interference of local patterns of action (reflexes) with the total action pattern of swimming would thus be prevented, and interference of the axial musculature of the right side with that of the left and vice versa, would be blocked."

Coghill's interpretations are in harmony with my own findings based upon the excision of these cells (Detwiler, 1927, 1933), where it was shown that the swimming reflexes of larvae lacking one or both neurones are inferior to those of normal animals. Larvae devoid of one Mauthner's neurone were found to exhaust more quickly than normals following repeated tactile stimulations. The swimming frequently was jerky and uncoordinated. There was no evidence from these experiments that the lack of one of the Mauthner's neurones had any effect upon equilibration.

In excising a Mauthner's cell, the ear vesicle first had to be removed and subsequently replaced by a vesicle from a donor embryo. Histological studies showed that, when larvae lacking the Mauthner's cell exhibited equilibratory disturbances, they were referable to an abnormal ear or to defective nervous connections, or both. It became apparent, from the results, that Mauthner's neurones are not indispensable to normal equilibrium. On the other hand, the absence of one ear was followed by marked equilibratory disturbances.

According to Coghill's view, the removal of Mauthner's fiber would release a normally present inhibitory effect on the local patterns of action and the interference between this and the total action pattern would be increased, as well as that between the axial musculature of both sides of the body. Regardless of the exact way in which the mechanism works, both the histological and physiological findings agree that this system is important in sustaining coordinated motor activity.

When the earlier experiments were made upon Mauthner's cells, no device was available for measuring the motor responses. It seemed desirable, therefore, to repeat the experiments and test their locomotor responses with the device shown in FIGURE 7.

The technique employed in removing the Mauthner cell consisted in first excising the ear vesicle. A small mass of cells was then cut away from the lateral wall of the medulla just medial to the ear region. This was followed by replacement of an ear vesicle from a donor embryo, using care that it was implanted with normal orientation. The operation was made on embryos in the early tail-bud stage (stages 27 and 28). For studies on comparative behavior, the ear vesicle alone was removed from a group of similar aged embryos, leaving the Mauthner neurone intact. In both groups, the *C* and *S* flexures and the early swimming responses were carried out in normal fashion. Beginning with stage 38, the motor responses were measured upon three groups as follows: (*1*) 10 normal control larvae; (*2*) 20 larvae from which the right Mauthner's cell had presumably been removed; and (*3*) 10 larvae lacking the right ear vesicle. This was done with the device shown in FIGURE 7. Each animal was stimulated 25 times, at approximately 5-second intervals, by stroking the trunk skin with a human hair as described above. The total number of units (sectors of arc) traveled by each animal following 25 stimulations was then computed, the average determined for the individuals in each group, and the standard error computed.* This procedure was carried out for each successive stage up to the feeding stage (44+) and then upon larvae 2, 4, and 7 days respectively after feeding.†

---

* I am indebted to Professor Herbert Elftman for subjecting my data to a statistical analysis and computing the standard error.

† In *A. punctatum*, the larvae begin to feed at stage 46. In these experiments, *A. jeffersonianum* was used and they were observed to begin feeding at a developmental period comparable in many respects to stage 44 of *A. punctatum*.

The data, based upon approximately 10,000 individual recordings, are given in FIGURE 24. An examination of this figure reveals a number of

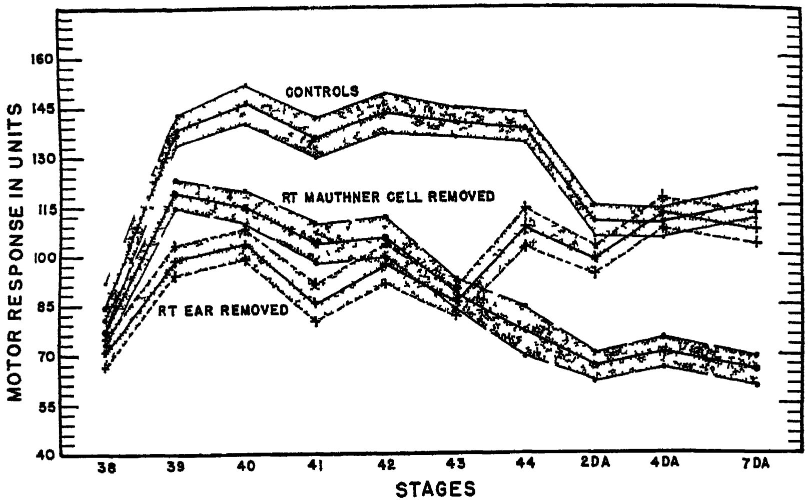


FIGURE 24 Graph showing average total distance traveled by larvae of *A. jeffersonianum* in response to 25 successive stimulations at 5-second intervals under the conditions indicated The heavy central line for each curve represents the average, the line above and below represents the standard error. The curves are based upon the data obtained from 10 normal larvae, 20 with the right Mauthner's cell removed, and 10 with the right ear removed Upon later histological examination, it was found that in 7 of the Mauthner neurone series the cell had not been excised (see text for explanation)

interesting facts. The larvae of all three groups made relatively low scores at stage 38, those lacking the ear being the lowest. At stage 39, there was a rise in the scores of all groups, but the controls markedly outdistanced those of the other two groups. Here, again, those lacking the ear made the poorest records; those lacking Mauthner's neurone occupied an intermediate position. From stage 40 on, there was a general overall decline in the motor ability of the group with theoretical absence of the Mauthner's neurone. Although there existed considerable variation at different stages, those lacking the ear showed marked improvement in their scores after stage 44, and at 4 days after the feeding stage their scores equaled those of the controls. Despite the improvement in their locomotor ability, these larvae lacking the ear still exhibited the characteristic equilibratory disturbances during locomotion. The behavior of larvae lacking an ear vesicle has been described by Greene and Laurens (1923) and subsequently by the author (1927, 1933). Such animals exhibit torsion towards the earless side. There is flexion of the arm on that side, with extension of the opposite arm. In swimming, the larvae rotate frequently on the longitudinal axis so as to exhibit a "corkscrew" movement. They may come to rest upside down, but more frequently land on the earless side. Spontaneous movements are usually less marked.

The larvae with theoretical absence of Mauthner's neurone showed no signs of improvement in their scores up to 7 days *post* feeding. Incomplete records of some older larvae of this group showed improvement in locomotor response, but the data were too scattered to be treated statistically. Viewing the curve as a whole, it is apparent that larvae lacking Mauthner's cell do not have the locomotor capacity of normal larvae. The fall in locomotor ability of this group began at stage 39. It may be pertinent to point out from microscopic studies of normal larvae that, at stage 39, Mauthner's cell shows some differentiation, which appears to be complete by stage 41.* According to the records (FIGURE 24), it is at this period when the motor response of this group began to decline (*cf.* controls).

In order to obtain evidence as to whether the Mauthner neurone had actually been excised, the larvae of this group were studied in serial transverse section. Whereas the group was originally made up of 20 larvae, only 16 were available at the end of the experiment. Both Mauthner's neurones were found to be present in 7 of the 16 animals. Most of these individuals, from which the right neurone had not been excised, also had low scores, but the average, on the whole, was higher than that of the group lacking this neurone. The low scores of the group with this neurone intact, as compared with those of the controls, leads one to suspect that, whereas the cell had not been excised, the operation may have disturbed the relations sufficiently to prevent normal connections of the cell body, resulting thereby in a lack of normal function. This theoretical explanation received some support from the fact that, in several cases, the perikaryon was displaced and in 2 cases was very small.

This series of experiments is regarded as incomplete, and it is planned to repeat the experiment and test the locomotor responses of much older larvae along with those of normal controls of similar age. In so far as they go, the results bear out former conclusions that the locomotor ability of larvae lacking the Mauthner fiber is inferior to that of normal animals.

Summary and Conclusions

1. With a device described in the text (FIGURE 7), the locomotor responses of young *Amblystoma* larvae have been quantified under the following conditions:

(*a*) Control larvae for each group listed below.

(*b*) Larvae lacking the cerebral hemispheres and the dorsal portion of the diencephalon.

(*c*) Larvae with the midbrain replaced by the anterior end of the spinal cord.

(*d*) Larvae with end-to-end reversal of the mesencephalon.

(*e*) Larvae with unilateral reversal of the mesencephalon (reversal of A–P axis only).

---

* Based upon a microscopic study of Harrison's normal stages

f) Larvae lacking the right Mauthner's neurone, but with intact ears.
g) Larvae lacking the right ear.

B. The data obtained from this group, where locomotor responses were measured from stages 39 to 46+, show that during these stages the motor capacity is only slightly lowered (FIGURES 16 and 17). The relative unimportance of the hemispheres in the general motor activities is clearly indicated.

C. Larvae without the midbrain, but with nervous continuity between fore- and hindbrain exhibit normal locomotor responses up to approximately stage 40. Thereafter, the motor capacity is greatly lowered (FIGURES 9 and 11). The data indicate that the swimming mechanism in the early stages is essentially spinal and autonomous, but that at approximately stages 40–41 it becomes subservient to mesencephalic control, probably concomitant with the development of the tecto-bulbar and tecto-spinal tracts.

D. Reversal of the mesencephalon only slightly lowers the locomotor capacity. The scores of larvae ranging from stages 39 to 46+ are in general of a magnitude similar to those exhibited by larvae lacking the hemispheres (*cf.* FIGURES 13 and 16).

E. Unilateral reversal of the midbrain (A–P axis only) has no effect upon the motor ability of the animal.

F. Animals lacking one Mauthner's neurone show significantly lowered scores as compared with those of control larvae (FIGURE 24). They differ from those lacking an ear in that the scores did not improve in the later stages under investigation.

G. The lowered motor ability of larvae lacking one ear is rather striking up to near the feeding stage. Thereafter, there is a gradual improvement. Four days after the feeding stage their scores equaled those of normal larvae (FIGURE 24).

The data obtained upon all groups are listed in graph form in the text. The method of study, as presented, gives an objective and quantitative picture of the performance of young larvae under various experimental conditions. It makes possible statistical validation of small differences which might be in doubt if reliance were placed solely upon subjective impressions.

Bibliography

BARTELMEZ, G. W. 1915. Mauthner's cell and the nucleus motorius tegmenti. J. Comp. Neurol. 25: 87–128.

BURR, H. S. 1916a. The effect of the removal of the nasal placodes on *Amblystoma* embryos. J. Exp. Zool. 20: 27–57.

1916b. Regeneration in the brain of *Amblystoma*. J. Comp. Neurol. 26: 203–211.

1930. Hyperplasia in the brain of *Amblystoma*. J. Exp. Zool. 55: 171–191.

COGHILL, G. E. 1909. The reaction to tactile stimuli and the development of swimming movement in embryos of *Diemyctylus torosus* Eschschultz. J. Comp. Neurol. & Psychol. 19: 242–261.

1934. New anatomical relations and the probable function of Mauthn [illegible] Psych. en Neuralog. Bladen 38: 386.

Detwiler, S. R. 1927. Experimental studies on Mauthner's cell in *Amblystoma*. Exp. Zool. 48: 15–30.

1933. Further experiments upon the extirpation of Mauthner's neurones in amphibian embryos (*Amblystoma mexicanum*). J. Exp. Zool. 64: 415–431.

1936. *Neuroembryology. An Experimental Study.* Macmillan. New York.

1944. Restitution of the medulla following unilateral excision in the embryo. J. Exp. Zool. 96: 129–142.

1945. The results of unilateral and bilateral extirpation of the forebrain of *Amblystoma*. J. Exp. Zool. 100: 103–117.

1946a. Experiments upon the midbrain of *Amblystoma* embryos. Am. J. Anat. 78: 115–138.

1946b. Midbrain regeneration in *Amblystoma*. Anat. Rec. 94: 229–238.

Detwiler, S. R., & W. M. Copenhaver. 1940. The growth and pigmentary responses of eyeless *Amblystoma* embryos reared in light and in darkness. Anat. Rec. 76: 241–257.

Greene, W. F., & H. Laurens. 1923. The effect of extirpation of the embryonic ear and eye on equilibration in *Amblystoma punctatum*. Am. J. Physiol. 64: 120–143.

Hamburger, V. 1946. Isolation of brachial segments of the spinal cord of the chick embryo by means of tantalum foil blocks. J. Exp. Zool. 103: 113–142.

Herrick, C. J. 1914. The medulla oblongata of larval *Amblystoma*. J. Comp. Neurol. 24: 343–427.

1939. Internal structure of the thalamus and midbrain of early feeding larvae of *Amblystoma*. J. Comp. Neurol. 70: 89–135.

Nicholas, J. S. 1930. The effects of the separation of the medulla and spinal cord from the cerebral mechanism by the extirpation of the embryonic mesencephalon. J. Exp. Zool. 55: 1–22.

Sharrer, E. 1932. Experiments on the function of the lateral-line organs in the larvae of *Amblystoma punctatum*. J. Exp. Zool. 61: 109–114.

FUNCTIONAL POLARIZATION IN DEVELOPING AND REGENERATING RETINAE OF TRANSPLANTED EYES*

By L. S. STONE

Department of Anatomy, Yale University School of Medicine, and Osborn Zoological Laboratory, Yale University, New Haven, Connecticut

STUDIES of the organization and location of the eye-forming centers in the brain wall (Adelmann, 1934, 1936, 1937; Alderman, 1935: Mangold, 1931; Stone and Dinnean, 1943) and the subsequent growth of those centers related to the induction and organization of the lens as shown by Spemann (1938) and many others, reveal how well this sense organ is adapted in amphibians as a tool for approaching problems dealing with polarization during development and differentiation.

The influence of the developing eye cup upon the formation of the medio-lateral axis of the lens has been intimated by the experiments of Woerdeman (1934) and LeCron (1907), but little notice has really been given to it. The outer pole of the developing lens forms the subcapsular epithelium. The inner pole is devoted to fiber formation. The evidence offered by LeCron and Woerdeman seems to indicate that the establishment of this axis may be normally dependent upon the length of time the lens anlage remains in contact with the optic vesicle and cup. When the lens placode was isolated early from the optic vesicle, it formed later only an epithelial vesicle. If the lens anlage was allowed to be associated with an optic cup for a period of time before it was isolated, it could later develop a fiber-forming pole. Even in some of these cases, the fibers would degenerate later. The optic cup is probably exerting a strong influence upon the establishment of this axis. There is a great need for further study of this problem.

Other organization within the lens can also be studied. Early on the anterior (distal) surface of the lens is a vertical suture and on the posterior (proximal) surface is a horizontal one. Focusing attention on the posterior (horizontal) suture, Woerdeman (1934) found that, by rotating the presumptive lens ectoderm 90°, this suture was in a labile form before the period when the neural folds became elevated, but it was determined just before the neural folds closed. Therefore, the polarity or growth direction of the fibers was already predestined at an early stage. According to Woerdeman, this is also borne out by the fact that, when the optic vesicle is rotated 90°, the suture line in the lens is unrotated, although the choroid fissure was rotated.

---

* Aided by grants from the John and Mary R. Markle Foundation and the Fluid Research Fund of Yale University School of Medicine.

Since the direction of the choroid fissure normally coincides with that of the anterior vertical suture of the lens, and is at the same time perpendicular to the posterior horizontal suture, this relationship could be used in further fruitful studies to uncover any influence the eye-forming centers may be exerting on the establishment of the lens sutures.

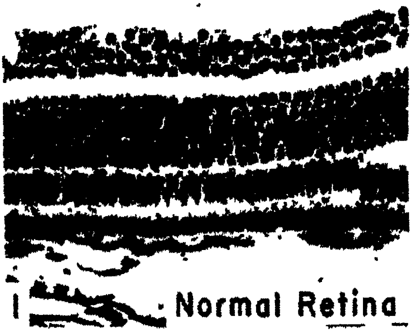
From the work of Beckwith (1927), Sato (1933), and Woerdeman (1934), one would judge that the time when the choroid fissure itself becomes determined varies in different species of amphibians. Since it can be used as a mark of polarization taking place in the eye cup, and since it marks the groove along which the early *lentis vasculosis* is directed in the higher vertebrates, it is highly desirable that the choroid fissure be examined even more closely than it has been in previous studies.

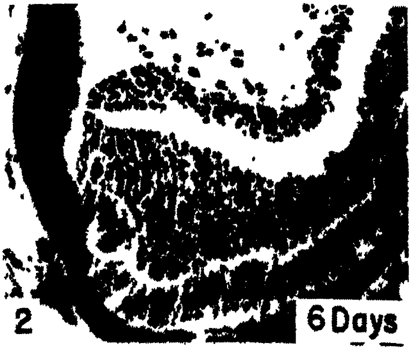
In the midst of all of these influences for organization, it may occur to the curious-minded that the future functional arrangement in the retina might also be manifesting itself. Therefore, as an interesting challenge from the standpoint of development and regeneration, we shall examine what evidence we have of this example of polarization taking place in the salamander eye.

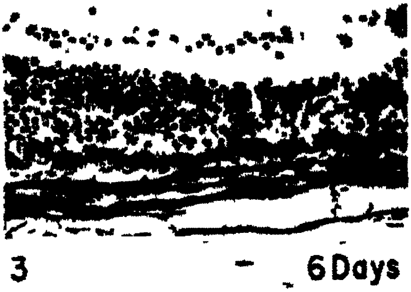
I have worked for many years on vision in salamanders, where the eye can be successfully transplanted followed by return of vision in the graft. Up to the present time, this has not been done successfully in other vertebrates. We have shown not only that the eyes of larvae (Stone, Ussher, and Beers, 1937; Stone and Cole, 1931 and 1943; Stone and Zaur, 1940: Stone, 1940) and adults can be transplanted to new hosts, and eventually vision returns in the graft, but that the eyes of quite different species of salamanders can be exchanged with equal success (Stone, 1930: Stone and Ellison, 1940 and 1945). This, I might say in passing, has presented an interesting method of studying visual acuity when it is different in two hosts whose eyes are being exchanged. Return of vision can be shown at least four times in the same adult salamander eye (*Triturus viridescens*) repeatedly transplanted to new hosts (Stone and Farthing, 1942). The adult salamander eye can even be successfully transplanted after seven days of refrigeration (Stone, 1946).

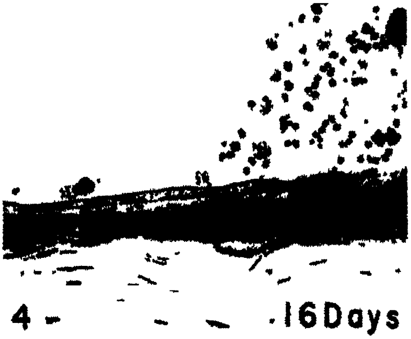
When the eyes are transplanted in larval salamanders, very little structural changes take place in the graft (Stone, Ussher, and Beers, 1937: Stone, 1930). The original retina is retained with an occasional loss of a few ganglion cells. The proximal stump of the optic nerve regenerates readily and follows along the pathway of the degenerated distal segment leading into the brain.

In the adult grafted eye of the salamander, a striking difference occurs. Not only all of the optic nerve degenerates but the original retina of the eye also disintegrates quite rapidly, with the exception of a ring of cells at the peripheral or ciliary margin (FIGURES 1–6). From these surviving cells, a new retina is regenerated and, after differentiation takes place, a new optic nerve grows back through the chiasma into the brain, where it apparently makes proper connections. Return of vision is usu









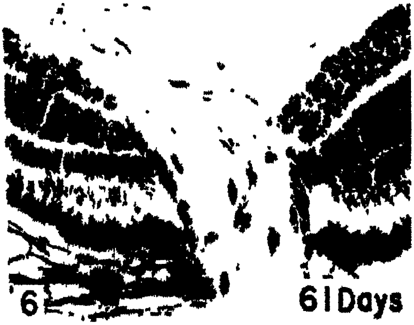


FIGURE 1. Photomicrograph of central portion of normal retina in adult eye of *Triturus viridescens*. x125 All other figures are from transplanted adult eyes of the same species.

FIGURE 2. Ciliary region in a transplanted eye 6 days after operation, showing far less degeneration than in the central retina (FIGURE 3). x125.

FIGURE 3. Same case as FIGURE 2, showing extensive degeneration in the central retina 6 days after operation. x125.

FIGURE 4. Showing complete degeneration of the central retina 16 days after operation. x125.

FIGURE 5. Showing a regenerating retina, three to four cells deep, in the central portion 23 days after operation. x125.

FIGURE 6. Showing the completely regenerated retina and optic nerve 61 days after operation. x125.

ally demonstrated in the transplanted eye between two and a half and three months after operation.

When Sperry, in 1942, first reported reversed visuomotor responses in adult salamanders after optic nerve regeneration in eyes which were rotated *in situ* without removal from the orbit, I examined the effects of various types of rotation on vision in many grafted eyes of *Triturus viridescens* in which the regenerated and new functioning retina could be tested (Stone, 1944). We shall return to these experiments later.

Since we shall be examining the functional quadrants of the retina, we need to know the normal visuomotor responses in the salamander, which are both simple and dependable for our tests. They are demonstrated when the animal moves toward, snaps at, and follows a lure approaching any of the four quadrants in the field of vision. The animal also automatically moves head and body in the same direction with a rotating black and white striped drum only if it passes from the temporal to the nasal poles (postero-anteriorly) through the field of vision of one eye (clockwise for a left eye and counterclockwise for a right eye). These compensatory movements are not elicited when the drum moves in the nasotemporal or antero-posterior direction through the field of vision of either eye.

In attempting to determine the stage in development at which the retina becomes functionally polarized, the right eye was excised, rotated 180°, and reimplanted in embryos of *Amblystoma punctatum* from the closure of the neural folds to a period just before the feeding stage begins. In other words, all the Harrison stages from 20 to 44 inclusive were examined. Vision tests during the larval and adult life of these hosts showed that, up to the late tail-bud stages, the eye could be rotated 180° without normal vision being affected later.

If the eye is rotated around the stage at which the first motor responses are known to appear in the embryo, the visuomotor reactions in the animal later on appear to be confused. Reactions sometimes appear slightly reversed but are not sharp and consistent. It is difficult to analyze them. However, if the eye is rotated at about the time when the beating of the heart is prominent, around Harrison stage 36, a definite reversal in the visuomotor responses will later become evident. The eye is now developed into a well-defined, broad, smooth-rimmed cup. A thin tapetum covers the outer surface of the thick undifferentiated future retina. The lens vesicle is just separating from the source of origin and will soon produce a fiber-forming pole. It fills the cavity of the optic cup against which it has been very tightly pressed ever since it sank inward as a lens plug. Anyone who has tried to remove the living lens anlagen at this time will know full well the implications of these remarks as applied to the operation. If the degree of adhesion with which they stick together is an indication of the importance of events now happening between then before they later separate, it must be all-powerful. In fixed preparations this relationship is poorly indicated as a distinctive feature. I believe it is worthy of far more attention than has been placed upon it.

Leading up to this period in development, the functional polarization of the retina is becoming established. Not long after this critical period, all rotated eye cups show that later in larval and adult hosts both the dorsoventral and antero-posterior (nasotemporal) axes are fully expressed by a complete reversal of the vision responses. Whether both axes are established at the same time or at different moments is being investigated by experiments now in progress.

In an earlier part of this discussion, attention was called to the fact that, as soon as one temporarily interferes with the blood supply to the retina in an adult salamander eye, such as transplantation accomplishes, the retina degenerates with the exception of a ring of cells at the peripheral or ciliary margin. These surviving cells aid in the regeneration of a retina which gives off a new optic nerve. This grows out to make connections with the brain, so that return of vision can be demonstrated between two and three months after operation. This offers an unusual opportunity not only to examine the possibility of return of vision but to test whether or not the functional quadrants in the retina can be re-established through the processes of regeneration.

In my studies on many hundreds of transplanted eyes in salamanders, it has been shown that normal vision eventually returns to the graft if it is normally oriented when placed in the orbit. In some manner, the new fibers from the ganglion cells in each of the quadrants in the regenerated retina in adult transplanted eyes register their stimuli with the proper centers in the optic tectum. It is difficult to conceive how normal vision could be re-established if this were not so.

To examine further the functional quadrants of the retina in the adult eye of *Triturus viridescens*, for example, we must perform a few simple operations involving various types of rotation experiments. In the first of these experiments, we shall rotate the right eye 180° anterodorsally *in situ* and fix it in place after cutting only the conjunctival and muscular attachments. Care is taken to insure that the optic nerve and blood supply are left intact in order to preserve the original retina. To prepare our animal for special tests, we shall remove the opposite left, normal eye. Following our usual technique, the animal, remaining under chloretone anesthesia, is placed in a cool moist chamber where it continues quiet for 24 hours. During this time, the right rotated eye becomes fixed and healed in place. Animals prepared in this manner are now placed in water in finger-bowl aquaria where they are kept for daily observations. We shall see that the eyes now possess all the quadrants of the original intact retinae functionally reversed.

The animals have a tendency to swim and walk in circles, sometimes in a very excited manner in short circles in the center of the aquaria with the head touching the tail. Their circuitous progression is almost entirely toward the blind side and, as they come to rest, the head continues to drift for a short distance in the direction in which they have been moving before it is brought back to the midline. When *Daphnia* are swimming about in the aquaria, these salamanders make erroneous strikes at the moving objects. Unless they make contact with their source of food, they may have great difficulty in obtaining it.

In response to a rotating black and white striped drum, the head and body movements are called forth only when the drum passes through the field of vision from the original temporal (posterior) pole to the original nasal (anterior) pole. Since the right eye has been completely rotated 180° in this case, the drum must be moving in the clockwise direction—the one which fails to call forth a reaction if the right eye is normally oriented. Instead of following in the direction of the moving drum, however, the animal walks or swims in the opposite or reverse direction. At first, the head starts drifting in the opposite direction. After reorientation, this nystagmoid movement may be repeated several times. Very soon the animal starts swimming and walking as if pursuing the drum, but always in the reverse direction. When the drum is rotated in the opposite direction (counterclockwise) through the field of vision from the original nasal (anterior) pole to the original temporal (posterior) pole, the animal usually assumes the position of a fixed stare. There is no nystagmoid movement of the head or body, and the animal makes no movements in response to the drum.

A dark object, such as a small piece of red rubber impaled on the end of a white wire, serves as a good lure for further vision tests. When the lure is brought from in front into the field of vision, the animal immediately seeks it in the opposite direction. When the lure approaches from the rear, the animal moves forward in its search. When the lure is brought above the animal, it immediately darts to the bottom of the aquarium to find the moving object, and when the lure is moved below the glass bottom of the aquarium the animal comes to the top of the water in its pursuit. In other words, all responses are completely reversed. If the animal is anesthetized again and the eye rotated back through the same arc to normal orientation, the swimming reactions and all visuomotor responses to the lure and drum are fully restored to normal as soon as the animal recovers from the anesthetic. We now have sufficient control observations to compare with the results obtained from rotated transplanted eyes with return of vision in a regenerated retina that replaces the original one.

Our next experiment, then, will be to excise a right eye, *for example*, and then reimplant it upside down. We have then rotated all retinal quadrants of the eye 180°, just as we did in the above experiment when the eye was rotated 180° *in situ* without removing it from the orbit. The general technique of operation is the same as that which I have often described before in the literature already cited. Since the degeneration and regeneration of the retina and optic nerve and return of vision will be completed at the end of two or three months, we must ex[illegible] the left normal eye before that time, in order to make our tests comparable to the one above, that is, exclusively on the rotated eye. We t[illegible] that, as soon as vision has returned, all of the visuomotor responses are as completely reversed as in the case where the eye was rotated in the orbit without destroying its original retina. If the eye is now rotated 180° in situ so that its axes are normally oriented, this places the retinal quadrants in normal position, and the visuomotor responses are normal again. Also, if a reimplanted rotated (180°) eye, which has been functioning for a considerable period with reversed vision, is transplanted a second time but normally oriented, the swimming reactions of the animal and the visuomotor responses are perfectly normal in every respect when vision returns again in the retina which regenerates for the second time.

We can study, in the same eye, the effects of rotation on two quadrants of the retina while the other two remain normally oriented (Stone, 1944; Sperry, 1945). This is done in two simple experiments, by excising either the right or left eye and grafting it to the opposite side. If, for example, we excise both eyes, discard the right one and, in its place, implant the left eye without inverting the dorsoventral axis, the quadrants in this axis are normally oriented but the antero-posterior (nasotemporal) axis is rotated 180°, thus reversing only the positions of the nasal (anterior) and temporal (posterior) quadrants. When vision returns in the new retina, the swimming, head movements, and reactions to a rotating drum are the same as when this axis (antero-posterior) was rotated in the eyes reimplanted upside down. Reactions to the lure were completely reversed and abnormal when the lure approached from in front or from the rear, but perfectly normal when it approached either dorsally or ventrally.

In another experiment, if we excise both eyes, discard the right one and, in its place, implant the left eye upside down this time, we maintain the normal orientation of the anterior and posterior quadrants but invert both the dorsal and ventral quadrants. Now, when the vision returns, if the lure is brought into vision above the water in the aquarium, the animal immediately seeks it at the bottom of the aquarium and *vice versa*. In seeking the lure moving into vision from the front or from the rear, the visuomotor responses were perfectly normal. Since this axis (antero-posterior) was not rotated, the head and body movements, swimming, and reactions to the rotating drum were the same as in animals with a single normally oriented right eye. Therefore, we can prove that the functional patterns in each of the retinal quadrants are re-established by regeneration and that the visuomotor responses are guided by the orientation which these retinal quadrants register to the central nervous system. How the peripheral and central connections are properly made to maintain these abnormal vision responses for over three years, is still a mystery and there is, as yet, no good evidence recorded in the literature to help us solve it.

There are other recorded studies of vision responses in amph following rotation of functional eyes. Some of the procedures of r tion have been carried out in a similar manner and others by a different from what I have already described. Sperry, in 1943, gave a d tailed account of the effects on vision after the adult eye of *Tritur viridescens* was rotated 180° *in situ*, preserving the original retina an optic nerve. Concerning abnormal swimming reactions, drifting he and body movements, the reversed reactions to the moving lure and th rotating drum, my observations corroborate his findings. To determine whether or not the regenerating optic nerve in a 180° rotated eye would connect with the brain in an orderly fashion to show the same abnormal pattern of vision, Sperry (1943a) rotated eyes 180° *in situ*, allowed them to heal in place, and then severed the optic nerve. In some, the cut ends of the nerve were twisted or crushed to increase all chances for confusion as the nerve fibers grew back to the chiasma and brain. In some cases, the blood supply to the retina was apparently not disturbed and the optic nerve must have regenerated without retinal degeneration such as Stone and Chace (1941) found. From his description of the gross appearance of some of the eyes and delay in return of vision, the blood supply to the retina must have been severed along with the optic nerve, in which case the retina was replaced after degeneration. In any event, all of his cases showed vision completely reversed, as in the case of those eyes which, as I mentioned, were reimplanted upside down.

Although one cannot transplant eyes of tadpoles or adult anurans successfully, because the retina will not regenerate (Stone, 1938 and 1940), the optic nerve will regenerate when cut if the blood supply to the retina remains intact. However, Sperry (1944) was able to extend his types of operations to larval and adult anurans by rotating the eye *in situ* and then cutting the optic nerve. The same effects on the visuomotor responses occurred in these amphibians as had been noticed in the salamanders.

Sperry also made another very interesting experiment. It is well known that, in amphibians, all the optic nerve fibers from an eye pass to the contralateral side. By destroying the optic chiasma and directing the proximal stump of the optic nerve to the pathway of the distal stump on the same side, Sperry (1945) forced the regenerating optic fibers in adult anurans to enter, on the ipsolateral side, the optic tectum with which it never had been and never is normally connected. Peculiar circus movements and mixed reverse reactions and rotating movements of the head and body, accompanied by ocular nystagmus, resulted after vision returned. The animals made errors in spatial localization when striking at a moving object or in trying to escape from an approaching large object. Further study is needed to understand these interesting reactions.

Our objective was to examine evidence dealing with the polarization of the functional quadrants of the retina. However, I think it has become obvious that the development of polarization within the optic tectum is a very intimate part of this general problem. To unravel it will be no small challenge to our ingenuity.

Summary and Conclusions

(*1*) Attention is called to the early polarization taking place in both the lens and the optic vesicle and to the interrelating influences being expressed between them at this time.

(*2*) 180° rotation of many stages of embryonic eyes in *Amblystoma punctatum* demonstrated that functional polarization of the retina is taking place in the optic cup of embryos shortly after the first motor responses appear.

(*3*) Function was also studied in the retinal quadrants as it was reestablished in regenerated retinae of rotated, grafted eyes of adult salamanders. In some cases, both the dorsoventral and nasotemporal (anteroposterior) axes were reversed (eyes reimplanted upside down). In others, one axis was reversed while the other remained normal (right and left eyes exchanged and properly rotated 180° on the one or the other axis).

(*4*) Complete reversal of visual perception occurred only in the rotated retinal quadrants, as exhibited by the visuomotor responses to a moving lure and to a rotating black and white striped drum.

(*5*) The reversed vision is permanent and only restored to normal when the eye is rotated back to normal orientation.

Literature Cited

Adelmann, B. H. 1934. A study of cyclopia in *Amblystoma punctatum* with special reference to mesoderm. J. Exp. Zool. 67: 217–281.

1936. The problem of cyclopia. Quart. Rev. Biol. 11: 161–182, 284–304.

1937. Experimental studies on the development of the eye. IV. The effect of the partial and complete excision of the prechordal substrate on the development of the eyes of *Amblystoma punctatum*. J. Exp. Zool. 75: 199–237.

Alderman, A. L. 1935. The determination of the eye in anuran, *Hyla regilla*. J. Exp. Zool. 70: 205–232.

Beckwith, C. J. 1927. The effect of the extirpation of the lens rudiment on the development of the eye in *Amblystoma punctatum*, with special reference to the choroid tissure. J. Exp. Zool. 49 (1): 217–259.

LeCron, W. L. 1907. Experiments on the origin and differentiation of the lens in *Amblystoma*. Am. J. Anat. 2: 245–257

Mangold, O. 1931. Determinationsproblem. III. Das Wirbeltierauge in der Entwicklung und Regeneration. Ergebn. Biol. 7: 193–403.

Sato, T. 1933. Über die Determination der fötalen Augenspalte bei *Triton taeniatus* Roux. Archiv. 128: 243–377.

Spemann, H. 1938. Embryonic development and induction. Silliman Lectures. Yale University Press.

Sperry, R. W. 1942. Reestablishment of visuomotor coordination by optic nerve regeneration. Abst., Anat. Rec. 84: (4): 20.

1943. Effect of 180 degree rotation of the retinal field on visuomotor coordination. J. Exp. Zool. 92 (3): 263–279.

1943. Visuomotor coordination in the newt (*Triturus viridescens*) after regeneration of the optic nerve. J. Comp. Neurol. 79 (1): 33–55.

1944. Optic nerve regeneration with return of vision in anurans. J. Neurophysiol. 7: 57–70.

1945. Restoration of vision after crossing of optic nerves and after contralateral transplantation of eye. J. Neurophysiol. 8: 15–28.

STONE, L. S. 1930. Heteroplastic transplantation of eyes between the larvae of two species of *Amblystoma*. J. Exp. Zool. 55: 193–261. (Harrison Anniversary Volume.)

1938. Return of vision and other observations in grafted vertebrate eyes. Am. J. Ophthalmol. 21 (1): 1–6.

1940. Reimplantation and transplantation of eyes in anuran larvae and *Fundulus heteroclitus*. Proc. Soc. Exp. Biol. & Med. 44: 639–641.

1941. Transplantation of the vertebrate eye and return of vision. Trans. N. Y. Acad. Sci. 3 (ser. II): 208–212.

1944. Functional polarization in retinal development and its reestablishment in regenerating retinae of grafted salamander eyes. Proc. Soc. Exp. Biol. & Med. 57: (1): 13–14.

1946. 'Return of vision in transplanted adult salamander eyes after seven days' refrigeration. Archiv. Ophthalmol. 35: 135–144.

STONE, L. S., & R. R. CHACE. 1941. Experimental studies on the regenerating lens and eye in adult *Triturus viridescens*. Anat. Rec. 79: 333–348.

STONE, L. S., & C. H. COLE. 1931. Grafting of larval and adult eyes in *Amblystoma punctatum*. Proc. Soc. Exp. Biol. & Med. 29: 176–178.

1943. Grafted eyes of young and old adult salamanders (*Amblystoma punctatum*) showing return of vision. Yale J. Biol. & Med. 15: 735–756.

STONE, L. S., & F. L. DINNEAN. 1943. Lens induction in the salamander (*Amblystoma punctatum*) with special reference to conditions in experimentally produced cyclopia. Yale J. Biol. & Med. 16: (1) 31–44.

STONE, L. S., & F. S. ELLISON. 1940. Exchange of eyes between adult hosts of *Amblystoma punctatum* and *Triturus viridescens*. Proc. Soc. Exp. Biol. & Med. 45 (1): 181–182.

1945. Return of vision in eyes exchanged between adult salamanders of different species. J. Exp. Zool. 100: 217–227.

STONE, L. S., & T. E. FARTHING. 1942. Return of vision four times in the same salamander eye (*Triturus viridescens*) repeatedly transplanted. J. Exp. Zool. 91 (2): 265–285.

STONE, L. S., N. T. USSHER, & D. N. BEERS. 1937. Reimplantation of larval eyes in the salamander (*Amblystoma punctatum*). J. Exp. Zool. 77 (1): 13–48.

STONE, L. S., & I. S. ZAUR. 1940. Reimplantation and transplantation of adult eyes in the salamander (*Triturus viridescens*) with return of vision. J. Exp. Zool. 85: 243–269.

WOERDEMAN, M. W. 1934. Über die Determination der Augenlinsenstruktur bei Amphibien. Z. Mikr.-Anat. Forsch. 36: 600–606.

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Volume XLIX, Art. 6. Pages 867–928
October 8, 1948

CURRENT TRENDS IN CLINICAL PSYCHOLOGY*

Consulting Editor: S. STANSFELD SARGENT

CONTENTS

	PAGE
Theoretical Foundations of Directive Psychotherapy. BY FREDERICK C. THORNE	869
Some Dynamic Aspects of Non-Directive Therapy. BY ARTHUR W. COMBS	878
The Theory and Practice of Group Psychotherapy. BY HELEN E. DURKIN	889
Psychodrama and Group Psychotherapy. BY J. L. MORENO	902
Discussion on Group Therapy	904
What Did the Clinical Psychologist Learn from the War? BY MAX L. HUTT	907
Future Training in Clinical Psychology. BY JAMES G. MILLER	913

---

* This series of papers is the result of a Conference on Current Trends in Clinical Psychology, held by the Section of Psychology of The New York Academy of Sciences on January 17 and 18, 1947. Publication made possible by a contribution from the General Funds of the Academy.

Copyright, 1948, by The New York Academy of Sciences

THE MERRYMOUNT PRESS, BOSTON, U. S. A.

THEORETICAL FOUNDATIONS OF DIRECTIVE PSYCHOTHERAPY

By FREDERICK C. THORNE
University of Vermont, Burlington, Vermont

Clinical Psychology in the Pre-War Period

IN the pre-war period, clinical psychology was in a state of extreme disorganization and individual uncoordinated effort. More than ever, clinical psychologists needed an approach to problems of diagnosis and therapy which was distinctively psychological and not merely a collection of second-hand theories borrowed from psychiatry, psychoanalysis and related disciplines. In the decades since Morton Prince first stimulated the clinical interests of American psychologists, an integrated and original system of clinical psychology had failed to evolve. Texts in the fields of abnormal and clinical psychology were rather sterile and did not contain enough which was distinctively psychological to establish upon it a sound foundation for clinical practice apart from the contributions of psychiatry and psychoanalysis. It therefore seemed desirable to re-evaluate existing theory and to formulate a genuinely eclectic system which, though not original or completely new, would at least integrate the potential contributions of experimental psychology and particularly the psychology of learning. War-time needs accelerated a revision and expansion of the concepts of directive therapy which had been envisaged before the war.

Non-Directive Methods. The decision quickly to publish a series of articles on directive therapy was precipitated by a peculiar situation which had developed in connection with the sudden recognition and popularity of the non-directive technique, as evolved by Rogers.[4] The impact of the new methods was felt in clinical psychology at a time when there was great war-time need for innovations which would fill the vacuum occasioned by the previous failure of psychology to develop clinical applications which were distinctively its own. As pointed out elsewhere,[8] Rogers is a very persuasive writer and his zealous enthusiasm caused the new methods to be accepted wholeheartedly and sometimes uncritically. There seemed to be a definite danger that the profession might abandon the hard-earned knowledge of the older directive methods in an impulsive adoption of the new. The danger seemed intensified by the opinions expressed by Rogers in his principal book,[4] to the effect that older directive methods were now outdated and discredited by the newer "client-centered" approaches. It therefore seemed desirable to make an up-to-date reformulation of directive methods in order to effectively counterbalance too uncritical an acceptance of the non-directive viewpoint.

In contrast with Rogers's non-directive therapy, which appears to represent a rather rigid technical system and which has attracted a group of enthusiastic followers who have formed what almost might be considered a new "school" of psychology, we do not believe that it is desirable or consistent with a genuine eclectic viewpoint to attempt to establish a complete school or system on the basis of the directiveness or non-directiveness of the methods used. In our opinion, directive psychotherapy does not involve a new school or system but merely a more objective evaluation of the methods and indications for utilizing direction in therapy. There appears to be definite value in objectively delineating the clinical relationship in a variety of situations, utilizing eclectically whatever method is indicated in a specific situation. Recognizing that both directive and non-directive methods are valuable in specific situations, the eclectic viewpoint denies the desirability of extremism in either direction. Ideally, a clinician should develop proficiency with all methods which have any valid application to any situation which may be encountered.

Terminological Difficulties. The use of such terms as directive or non-directive to describe systems of therapy is probably unfortunate. As used by Rogers, "non-directive therapy" refers to a specific and sharply delineated system in which extreme non-direction is the most significant characteristic. It is contended that non-directive therapy is a system complete in itself and mutually exclusive of the more orthodox directive methods. Anything less than complete non-directiveness is not considered compatible with being a genuine non-directive therapist. In practical usage, semantic difficulties have arisen from the attempt to apply an "either-or" dichotomy to what is actually a continuum. Rarely is therapy either completely directive or non-directive, and there is a question whether genuine indication commonly exists for extremism in either direction. Most workers are in agreement that non-direction is very valuable in certain situations but do not agree that it is the only and inevitable basis for sound therapy.

Further misunderstandings have developed over the differences, if any, between the "passive" methods developed in psychoanalysis and social case work and the allegedly new non-directive technique. Lowery[3] and others have emphasized that the basic principles of non-direction were known and in common use before the work of Rogers, whose contribution seems to have been in carrying them to further extremes than ever before. The principle of being "client-centered" has received general acceptance and is not the sole property of any school. Many of the criticisms leveled by Rogers[4] against the older directive methods are quixotic in the sense that they deal with undesirable practices abandoned long before by competent therapists everywhere.

In spite of semantic difficulties, the term "directive psychotherapy" was chosen to emphasize that direction does have a legitimate place in therapy at one end of a continuum of which non-directive methods stand at the opposite extreme. In our own personal practice, we attempt to be genuinely eclectic, utilizing whatever degree of directiveness seems indicated in each specific situation.

The Nature of Clinical Science

The most valid theoretical foundations for directive psychotherapy are derived from a survey of the historical development of clinical science. As pointed out elsewhere,[7] clinical psychology has a medico-psychological heritage dating back at least to the founding of modern medicine by Hippocrates over 2,300 years ago. The pre-scientific era of medical psychology, in which mental disorders were recognized and treated with primitive religious devotions, has been traced back at least to 5000 B.C. in ancient Egypt.[9] The significance of these facts is that one must be familiar with the history of medical psychology in order to evaluate recent developments in the proper perspective. There has gradually accumulated a knowledge of mental disorder and its treatment which, at least as practiced by medical physicians, has been taught by word of mouth and precept rather than objectively in books. This statement is true of all the medical specialties, *i.e.*, many details of hospital or surgical practice are not to be found in any book but must be learned in actual experience. This observation is particularly true in the specialty of psychiatry, in which most of the textbooks discuss general theory and omit the details of actual practice. Thus, it happens that there are not yet available objectively oriented, detailed statements concerning what happens in modern directive therapy. Rogers has admirably objectified the formal technique of non-directive therapy, but, as yet, comparable studies are not available for the more complex techniques of directive methods.

Methods of Clinical Science. Clinical science utilizes the same techniques of description, classification, statistical analysis, and explanation as are standard in experimental laboratory science. Operating upon the theoretical foundation of a thoroughgoing materialism, clinical science attempts to discover the causal principles involved in disease and then to discover, by experimental methods, specific treatments for each pathological condition. The fact that description and classification in clinical science have the additional objectives of diagnosis and therapy does not detract from their essential validity, since all steps in the process may be objectively quantified by the same methods of analysis as are applicable to any other type of data. The essential identity of objective description and diagnosis is illustrated by the fact that, the more comprehensive the description, the more complete the diagnosis. It should be clearly recognized that the foundations of modern clinical science rest upon adequate etiological studies which have the objective of describing, identifying, and diagnosing the causative factors of disease without which rational therapy is impossible.

Diagnosis in Clinical Science. If it is accepted that valid diagnosis is the foundation of rational practice, it follows that some degree of directiveness is present in every clinical relationship, since the clinician must conduct certain examinations to describe the status of the patient. Rogers[5] has recently denied that diagnosis is either desirable or necessary as a foundation for non-directive therapy, on the grounds that even this amount of directiveness contradicts the basic principle of being "client-centered." Such a viewpoint is in complete opposition to the accepted principles of modern clinical science, which state that valid diagnosis is fundamental to rational therapy. It would appear that even the most completely non-directive technique involves at least two diagnostic decisions, *i.e.*, (*a*) that the etiological cause of the disorder involves emotional problems which are amenable to non-directive therapy, and (*b*) that the client has sufficient personality resources to resolve his own problems without directive support. It must be clearly recognized that "blind" or "shot-gun" treatments have long been discredited in modern clinical science even though they may be effective in many minor conditions in which any suggestive method would produce some alleviation of symptoms. The fact that it is possible to treat some relatively superficial conditions by non-directive methods and without adequate diagnostic studies does not necessarily prove that these methods are either scientific or valid in other more serious disorders.

Objectives of Diagnosis. In modern clinical science, the objective of diagnosis involves more than identifying and naming a pathological syndrome. Recognizing that the personality dynamics in each individual case are different, psychodiagnostics has moved beyond problems of classification to the more mature objective of completely describing the etiological factors causative of disorder. Among the objectives of psychodiagnostics are:

(*1*) To demonstrate the etiological factors.

(*2*) To differentiate between organic and functional disorders.

(*3*) To discover the personality reaction of the organism to the disorder.

(*4*) To evaluate the degree of organic and functional disability.

(*5*) To estimate the extensity or intensity of the morbid process in relation to actuarial data concerning type and severity.

(*6*) To determine a prognosis or probable course.

(*7*) To provide a rational basis for specific psychotherapy.

(*8*) To provide a rational basis for discussing the case with patient and relatives.

(*9*) To provide a scientific basis for classification and statistical analysis of data.

(*10*) To formulate a dynamic hypothesis concerning the nature of the pathological process, and the mechanisms whereby therapeutic effects are explained.

Accumulated experience with problems of diagnosis in medical science has resulted in the discovery of a number of basic principles for diagnosis which are presented elsewhere.[7] Until such time as orthodox methods of clinical practice as developed in clinical science in general are proven to be invalid for clinical psychology, it is our opinion that the clinician has a responsibility for carrying out comprehensive diagnostic studies involving whatever degree of directiveness may be necessary to obtain the desired facts.

Physician-Patient Relationships. A distinctive psychosocial relationship holds in all the healing arts in that here is a situation where one individual expects and demands that another individual shall be more intelligent and have more training and experience. The patient seeks the help of the physician because he has problems which he is unable to solve with his own resources, *i.e.*, he turns to someone who is presumably more capable of achieving a solution. This dominance-submission relationship is present whether the therapy is directive or non-directive. Even with the most extreme non-direction, the patient is submissive in that he seeks treatment from another who presumably dominates the situation even though his method requires that as much responsibility as possible be turned back to the patient. The insistence that non-directive therapy be completely "client-centered" does not alter the fact that the therapist determines what is done and must therefore accept responsibility for the results.

In our opinion, the critical factor in determining whether therapy should be directive or non-directive relates to the decision as to whether the principle of homeostasis can be depended upon to produce a solution of the client's problems. Rogers[5] emphasizes that non-directive therapy seeks to provide a situation in which the potential personality resources of the client are utilized to the utmost in achieving a solution. Such an assumption may be tenable with minor personality disorders in which the client retains relatively good personality integration and is well enough to regulate himself without direction from without. In more serious personality disorders, however, the client may not possess the resources to achieve adjustment through normal homeostatic processes. In such cases, it is the duty and responsibility of the clinician to use such directive methods as may be necessary to supplement the client's resources.

Sutich[6] and others have somewhat confused the basic issues by identifying the non-directive method with so-called democratic principles. It is contended that the counselor has no right to undemocratically interfere with the personal integrity and rights of the client. In our opinion, this viewpoint is illogical and irrelevant since it ignores the basic fact that, the more mentally disordered a person, the more irresponsible he is, and the greater the necessity for regulation by some external agency until such time as the client regains mental health and resumes self-regulation. One of the basic principles of psychoanalysis and other forms of psychotherapy is that treatment is not complete until the patient has been returned to independent living. Clinical practice must be conducted upon the ethical principles of the profession rather than upon ideological systems of political origin.

Legal Responsibilities. Although the matter has not yet come to legal test in the field of clinical psychology because of the relative youth of the specialty, it seems to follow that, as a practitioner of a healing art, the clinical psychologist must be held legally responsible for the ethical and competent practice of his profession according to the accepted standards of time and place. The principles of malpractice in the longer established healing arts are now well set up and are probably applicable to clinical psychology without any basic modification. Having once been approached by a patient for consultation, and having accepted the relationship by entering upon consultation, the physician assumes legal responsibility for his actions and is liable to be sued for malpractice in the event that he is proven to be either civilly or criminally negligent. Errors of omission or commission caused by failure to observe ethical principles or to apply accepted methods of treatment make the physician legally responsible for results caused by such malpractice.

The nature of this legal responsibility probably implies that the physician shall be sufficiently directive to comply with requirements of accepted practice. There are many clinical situations in which the client must be protected from the consequences of his own actions, *e.g.*, when he has suicidal impulses. There are many other clinical situations in which the client does not have sufficient resources to solve his problem and must depend upon external support. Accepting the general premise that non-directiveness is the method of choice where there is high probability that it will be effective, there are a wide variety of clinical situations in which some degree of directiveness is indicated to protect the best interests of the client. From a legal viewpoint, it is assumed that the physician exerts directive control over the clinical relationship at all times unless it can be demonstrated that the patient voluntarily refused to cooperate or otherwise prevented a satisfactory outcome.

Part of the legal responsibility of the physician is concerned with the establishment of a valid diagnosis in order to discover or rule out the existence of malignant pathological processes which might seriously injure the physical or mental welfare of the patient. It is the physician's responsibility to search actively for pathological processes which, if undiscovered, might constitute a basis for the charge of malpractice. Particularly with clients who have seriously diminished personality resources and integration, it does not appear that non-directive methods can be depended upon to effect a diagnosis and adequate treatment.

Therapeutic Considerations

Although therapeutic practice will probably never become as scientifically objectified as is possible in the field of diagnosis, the accumulated experience of clinical science has as valid an application in therapy as in diagnosis. Defining therapy as including all types of formal case handling derived from scientific evaluation of the individual case by competent personnel, rational therapy (whether directive or non-directive) should proceed logically from etiological studies, clinical examinations, and laboratory studies from which a diagnostic formulation results. Depending upon therapeutic indications, methods of case handling may range from the most superficial contacts, as in counseling, to the most intensive depth therapy, as in psychoanalysis.

Modern therapy should be based upon a rigid adherence to materialistic concepts of disease.[1] It is necessary to have a thorough knowledge of gross and microscopic pathology, pathological physiology, and psychopathology in order to comprehend the etiology of the disorders of the total organism. Therapy which is not founded upon the most comprehensive possible understanding of psychopathology cannot be considered to be thoroughly modern and scientific. It is difficult to comprehend how therapy can fulfil the basic requirements of being (*a*) oriented on the basis of adequate diagnostic studies, (*b*) directed toward the correction of etiologic factors, and (*c*) executed on the basis of a detailed knowledge of the limitations of methods unless the therapist assumes responsibility for directing the basic outlines of treatment. Rogers' insistence ([5], p. 421) that diagnosis may be dispensed with and full responsibility for the conduct of treatment be placed upon the client is not consistent with the accumulated experience of almost all other clinical sciences.

Eclecticism in Therapy. Never in the history of clinical science has there been discovered a universal panacea, or method to end all methods, which is applicable to all types of morbid conditions. A study of the history of clinical science will reveal that the initial enthusiasm which greets all genuinely valuable new discoveries is followed by a period of more sober evaluation in which the indications and limitations of the new method are gradually determined. Thus, the sulfonamides and penicillin were originally believed to have been much more miraculous than later proved true. These comments also appear to apply to non-directive therapy, which is valuable but not the panacea which some have considered it to be.

The principle of eclecticism is the keynote of modern clinical science. To be eclectic is not to identify one's practice with the theories of any one system or "school" of thought. In our opinion, both directive and non-directive methods are valuable when utilized according to a comprehensive knowledge of their indications and contra-indications related to the needs of each specific case.

Directive Psychotherapy. The following outline presents the basic pattern of directive psychotherapy in which the therapist, though client-centered, assumes responsibility for conducting all details of case handling according to the highest ethical and professional standards of time and place.

1. *Adequate diagnostic studies,* involving
   a. Complete case history.
   b. Clinical examinations.
   c. Psychometric and projective studies.
   d. Laboratory procedures such as electroencephalography.
2. *Making a descriptive formulation* of the psychodynamics of each individual case, including etiology, clinical status, personality resources, and prognosis.
3. *Outlining an individual plan of therapy* with client-centered orientation which is specifically related to the needs of the individual case.
4. *Genuine eclecticism* in therapeutically utilizing all the technical resources, either directive or non-directive, which are available at time and place.
5. *The principles of experimental science* should be utilized wherever applicable at all levels of case handling, and particularly in etiological studies and psychodiagnosis.

In our opinion, this outline of directive psychotherapy is consistent with the historical evolution of clinical science in general and medical psychology in particular. It combines the best characteristics of experimentalism and modern clinical science.

Bibliography

1. Bellak, L., & R. Ekstein. 1946. The extension of basic scientific laws to psychoanalysis and to psychology. Psychoanal. Rev. 33(3).
2. Lecky, P. 1945. Self-Consistency: A Theory of Personality. Island Press. New York.
3. Lowrey, L. G. 1946. Counseling and therapy. Am. J. Orthopsychiat. 16.
4. Rogers, C. R. 1942. Counseling and Psychotherapy. Houghton Mifflin. Boston.
5. Rogers, C. R. 1946. Significant aspects of client-centered therapy. Am. Psychologist 1: 415–422.
6. Sutich, A. 1944. Toward a professional code for psychological consultants. J. Abnorm. Soc. Psychol. 39: 329–350.
7. Thorne, F. C. 1947. The clinical method in science. Am. Psychologist (In press).
8. Wrenn, C. G. 1946. Client-centered counseling. Educ. Psychol. Meas. 6: 439–444.
9. Zilboorg, G. 1941. A History of Medical Psychology. Norton. New York.
10. Andrews, J. S. 1945. Directive psychotherapy: I. Reassurance. J. Clin. Psychol. 1: 52–66.
11. Steinmetz, H. C. 1945. Directive psychotherapy: V. Measuring psychological understanding. J. Clin. Psychol. 1: 331–335.
12. Thorne, F. C. 1945. Directive psychotherapy: II. The theory of self-consistency. J. Clin. Psychol. 1: 155–162.
13. Thorne, F. C. 1945. III. The psychology of simple maladjustment. J. Clin. Psychol. 1: 228–240.
14. Thorne, F. C. 1945. IV. Therapeutic implications of the case history. J. Clin. Psychol. 1: 318–330.
15. Thorne, F. C. 1946. VI. The technique of psychological palliation. J. Clin. Psychol. 2: 68–79.

16. Thorne, F. C. 1946. VII. Imparting psychological information. J. Clin. Psychol. 2: 179–190.

17. Thorne, F. C. 1946. VIII. The psychology of satiation. J. Clin. Psychol. 2: 261–266.

18. Thorne, F. C. 1946. IX. Personality integration and self-regulation. J. Clin. Psychol. 2: 371–383.

19. Thorne, F. C. 1947. X. Constitutional analysis. J. Clin. Psychol. 3: 75–83.

20. Thorne, F. C. 1947. XI. Therapeutic use of induced conflicts. J. Clin. Psychol. (In press.)

SOME DYNAMIC ASPECTS OF NON-DIRECTIVE THERAPY

By ARTHUR W. COMBS
School of Education, Syracuse University, Syracuse, New York

ALTHOUGH it is only five years since the publication of Dr. Rogers'[4] book, *Counseling and Psychotherapy*, non-directive therapy has captured the attention and imagination of psychologists in a way that has seldom been accorded such an innovation before. At the present moment, it has become a lively and controversial issue in the psychological literature. This vigorous growth and development has been accompanied by inevitable changes in both theory and practice as new elements have been discovered and fitted into place.

Non-directive therapy originally grew out of the experience of many workers dealing with human adjustment and became established as a technique for the very concrete reason that it worked. Faced with the necessity for dealing with clients in everyday work, it is not surprising that the earliest efforts of non-directive therapists should have been turned to the improvement of their techniques and exploring its uses. As a result, the past five years have led us to consistent and well-established basic techniques for non-directive therapy although we are by no means so clear as to why and how these methods operate. We are finding out, however, and some of the dynamic and theoretical bases of non-directive therapy are proving to be no less exciting and far-reaching than the results of therapy itself. It is the purpose of this paper to discuss very briefly some of these current approaches to theory and practice.

The Operation of Need in Therapy. Non-directive therapy is based upon the fundamental principle that *the client not only can, but will, move toward better adjustment when an adequate situation is provided which frees him to do so.* Actually, the concept that the individual not only can, but will, move toward better adjustment is fundamental to therapy of any sort. What distinctly characterizes non-directive therapy in this connection is that it not only recognizes the existence of this drive, but consciously attempts to utilize it for therapeutic purposes.

Non-directive therapists, in discussing their philosophies and techniques, have repeatedly spoken of "the drive of the individual toward growth, health, and adjustment." After the wealth of cases which have now been published, something approaching this drive can certainly be called characteristic of the process. It may even be observed in many cases that the client moves toward a healthy condition even though such movement may be accompanied by the most extreme distress for himself.

Time after time, in the protocols of counseling, the client's own statements reveal his feeling that something approaching this drive is operating within him. For example, one client says at the moment of blackest despair in her counseling:

"I've made a bit of progress here. I'll make more as time goes on, I'm sure. You know, I think sometimes just to get it out of my system will help."

At the very end of the same interview, the client makes this remarkable statement:

"Well, the worst is out—next week we can start reconstructing!"

From the wealth of evidence already accumulated, there can be no doubt that non-directive therapy is often extremely, even dramatically, successful in aiding its clients in a wide variety of problems. When one considers that, in this type of therapy, the counselor has remained "non-directive" and has carefully refrained from the slightest hint of coercion or suggestion, it is not possible for such changes to have been brought about by the counselor. We must presume, therefore, that whatever this motivating force, it has its origin within the organism itself. This is consistent with much of our modern thinking about the nature of the organism. In biology, the drive to maintain self-organization has long been recognized and is described as fundamental to all living things, in the principle of "homeostasis." The very science of medicine has been built upon the concept that the organism can and will return to a state of health if the blocks to its recovery are removed. In view of our present knowledge of the unitary character of the organism, it would be extremely queer if this function were not true in psychological realms as well. We are by no means lacking in evidence that the same function does occur in the realm of psychological processes, for it is a well-known observation that many of our mental patients get well without outside assistance and even, sometimes, in spite of assistance. We might describe this need, as it operates in psychology, as a need to maintain or enhance the individual's personal organization.

One might raise the question at this point that, if the client has a drive toward health, if he can and will move toward adjustment, why is therapy ever necessary? The answer seems to be that sometimes the organism's need to maintain or enhance its personal organization is blocked by forces preventing its maximum fulfilment. Such blocks may arise, externally, in the individual's environment, or, internally, within the individual himself.* In either case it is likely to result in confusion and failure of accurate perception. This, in turn, results in random behavior, or what has been called "escape behavior," offering only a temporary solution to the individual's problems, and, more often than not, even complicating them further. By giving the client temporary relief and making it possible to protect his organization for a time, "escape behaviors" block more adequate perception in terms of which a more permanent solution might be achieved. Thus, confusion and distortion of perception play a large part in the production of the individual's maladjustment. Being unable to differentiate his situation clearly, the client may be unable to find an adequate solution and so continue his maladjustment indefinitely.

---

* It will be recognized that this is true in medicine as well, wherein disease, defined in its broadest sense, is a state in which normal bodily processes are obstructed by either external or internal attacks upon the organism's organization.

But another factor enters the situation which further prevents the free activity of the drive about which we have spoken. This is the effect of threat upon the individual and his self-organization. In addition to the confusion and failure of adequate perception in maladjustment, it will be recognized that one of the most commonly described characteristics of maladjustment, and particularly of neurosis, is the feeling of fear expressed in one form or another by the subject. Oftentimes, the object of fear is vague and undifferentiated. Such feelings of threat are characteristic of situations in which the individual is blocked from achieving the satisfaction of his fundamental need to maintain or enhance his self-organization. What is more, the closer such blocks come to affecting the organization of self, the greater the severity of the threat felt by the individual and the more active become his attempts to find a solution to his problem. Feelings of threat have unfortunate effects upon the client's ability to make a satisfactory adjustment.

We have stated previously that the basic need of the organism is to maintain or enhance its personal organization. When this basic organization is threatened, however, the organism has no choice but to defend the organization which exists, so that a movement toward a more adequate organization becomes impossible. Indeed, there is even danger that, if the threat is great enough, the client may be driven deeper into his present organization and his maladjustment made greater. This effect of threat has not been given much consideration in psychology although it is well known to the layman and the advertising expert or salesman. It is common observation, for instance, that "nobody ever wins an argument," and the quickest way to get a man excited is to threaten the things he holds most dear, usually himself. We may even hope that, some day, our diplomats will come to recognize this principle of human relationships. In our own science, we have long recognized the principle that "aggression yields aggression," but too often we have given mere lip service to the idea and in applied psychology have often acted as though it never existed.

We have seen in this theoretical discussion that the individual can and will move toward improved self-organization, but that this is often prevented by the client's own confusion or failure of adequate perception and feelings of threat which hamper or preclude movement toward new self-organization. It would appear, if this analysis is correct, that if we can remove an individual from threat and assist him to differentiate his personal organization more clearly, he should be free to move toward a more adequate and satisfactory self-organization and hence to better adjustment. This is precisely what non-directive therapy attempts to do. It consciously creates a relationship between therapist and client which scrupulously protects the client from threat, while, at the same time, encouraging and assisting him to a clearer and more accurate differentiation of his self-organization and its relationship to the world in which he moves.

Elimination of Threat. In his counseling relationship with his client, the therapist attempts to eliminate threat by an attitude of sympathetic understanding and acceptance and a careful avoidance of any action which might be construed by the client as a violation of his integrity.

In the early days of non-directive therapy, a great deal of stress was placed upon counselor techniques by means of which the individual was thought to be helped to his adjustment. Recently, however, there is a growing feeling among therapists using this approach that the atmosphere of the counseling relationship is of even greater importance than techniques. By this atmosphere is not meant the physical surroundings or the prestige of the counselor, but a "permissive atmosphere" consciously created by the counselor in his relationship with the client. In the therapy session, the client finds himself in a situation characterized by warmth and responsiveness on the part of his counselor, free from pressure and coercion of any sort, and in which he may express himself in any way he pleases within the very broad limits of the counseling relationship, which Rogers[4] describes as follows:

"From the client's point of view, while he may not be conscious of these elements at the outset, he does respond to the atmosphere of freedom from all moral approval or disapproval. He finds that he does not need his customary psychological defenses to justify his behavior. He finds neither blame nor over-sympathetic indulgence. He finds that the counselor gives him neither undue support nor unwelcome antagonism. Consequently, the client can, often for the first time in his life, be genuinely himself, dropping those defensive mechanisms and over-compensations which enable him to face the world in general."

As Rogers has suggested, the therapy situation is probably vastly different from any the client has ever before encountered. In daily life, the person who desires to tell others of his problems finds others all too ready to tell him theirs. Even worse, he is likely to be subjected to attempts on the part of those around him to force change on him in some fashion or other. This does not occur in non-directive therapy, however, and the client usually recognizes and makes use of the opportunity provided him. Note how the following clients express their understanding of this "special" kind of atmosphere created in the therapeutic relationship.

One client says:

> "When I'm talking to you, I just can't lie to you. Why, I can't even lie to myself. I don't think I've ever been so honest with myself before."

Another painfully shy young woman, extremely frightened by men, tells her male counselor:

> "In here I can talk to you. But if I were to meet you out in the hall you would be just another man to me and I couldn't talk to you at all."

The same client writes two years after her experience in counseling:

> "My counseling experience came at a time when my life was most disorganized. I went to you as a last resort. It was distasteful to talk to anyone about my problems—but the more I talked and the less you seemingly interfered, the better I felt and the more clearly I saw what the trouble was. Since that experience, my life has become more positive and creative. I have gradually placed more confidence in my own abilities and opinions. I feel completely happy and organized most of the time. I feel at ease with people of both sexes and all ages and have no trouble in meeting or liking the hundreds of people I have come in contact with on this new job."

In such a sheltered atmosphere, protected from the necessity of self-defense, the client finds himself free to examine himself in any way he desires. He finds no blocks to free expression placed in his path, and can utilize the therapy session in his own way as his needs determine.

It may be argued that no counselor worth his salt goes about threatening his clients. Experience has shown, however, that even the best-intentioned counselor cannot avoid threatening his client when he uses directive techniques. Even so simple an act as answering a knock at the door may be construed as threatening by the client, given the proper circumstances. This has actually happened in our experience. After answering the door and briefly directing the inquirer to another office down the hall, the counselor returned to his client. Note the feeling of threat implied in what the client said next:

> S.—"You know, I thought when you went to the door, you would tell me that's all for today."
>
> C.—"You thought I really wasn't interested enough to go on today, is that it?"
>
> S.—"That's right. I know it's silly, but I find I am very sensitive to anything you say or do. I'm so afraid you don't really want to help me. I'm afraid you might think the things I have to say are silly and I keep expecting you to tell me something that will upset me more, like the others I've been to—Oh! Don't misunderstand me!—I know by now that you won't, but I still get panicky inside when I think you might not want to help me."

It must be recalled that it is the client who feels threatened and not the counselor. The threat to organization lies not only in the actions or words of the counselor but in the peculiar meanings of these behaviors to the client. What is more, since these meanings exist within the client, they are not open to observation by the counselor but lie outside his control. Frequently, the client may feel threatened without giving the slightest sign of this fact to the counselor. Most clients have learned, through long experience, how to keep their feelings from being observed, but even if they had not, it is too late to deal with the effect of threat after it has occurred and the client has reacted. Severe threats to the client may even in some cases be sufficient to prevent him from ever attempting to examine a particular line of thought once he feels this is inappropriate or would be disapproved by his counselor. To avoid such threats consistently is by no means easy, as many a beginning counselor has discovered.

As we have suggested in a previous paper,[1] protection of the client "makes possible the pursuit of a question to its ultimate conclusion, a process greatly impeded by directive remarks." This freedom to examine and to pursue a line of thought to its "bitter end" makes insight possible with greater speed than is otherwise practicable. As one client put it: "It's like clearing away the brush that confuses the path." Freed from the necessity of defense and able to carry out such "bitter end" analyses, it becomes possible for the client to explore his field of meanings in any way he desires and to arrive, eventually, at a clearer differentiation of his self-organization and his relationship to the world in which he operates.

Aiding Differentiation. The second major function of non-directive therapy lies in assisting the client to more adequate differentiations. As we have seen, maladjusted clients characteristically are confused and have failed to differentiate clearly the various aspects of their self-organization and the environment in which they operate. Unable to differentiate clearly the various aspects of his field, the client in trouble finds himself anxious, worried, and vaguely afraid of something he cannot quite put his finger on. So long as this situation continues, a change in self-organization seems very unlikely. Indeed, it is likely to be very nearly impossible.

To assist his client to make such differentiations, the non-directive counselor does two things: (*1*) he encourages his client to express himself freely and in any way he pleases; and (*2*) he assists him to examine his personal meanings through a technique known as "recognition and acceptance of feeling." From a theoretical point of view, we might just as well call this technique recognition and acceptance of personal meaning, for that is essentially what it is. Feelings, after all, are simply the client's way of expressing the meaning of a situation for him. They express his personal reaction to the situation he is describing.

Personal meanings are crucial in the client's behavior, for we behave not in terms of events but in terms of the meanings of events for us. If Mrs. Jones, for example, feels that her husband is a brute and a heel, it makes little difference in her behavior whether he actually is or not. If she feels that way, it is enough. What is more, pointing out to Mrs. Jones that her husband is really not a brute but a very fine fellow will be of little help to her since this is not her opinion but the therapist's, and may only serve to demonstrate that the therapist "just doesn't understand." It will be recognized, too, that such a statement by the therapist constitutes a threat to Mrs. Jones and may force her to defend her position more tenaciously than ever. Telling the client is of little value if he cannot accept such information into his personal organization of meanings. It is the personal meaning of facts which motivate behavior, not the facts themselves.

It is the very fact that the individual's personal organization of meanings is inacceptable to others that classifies him as maladjusted. In adjustment counseling, therefore, some reorganization of these personal meanings must occur if therapy is to be truly effective and if any change in behavior is to result. Thus, the counselor in non-directive therapy assists his client to explore his personal meanings by recognizing and accepting those meanings as they appear. This helps his client to explore his personal organization of meanings further, helps to clarify them and holds them up for further examination. Under this kind of treatment, the client is able to make greater and greater differentiations in his meanings, resulting eventually in a new organization of personal meanings, as a result of which change in behavior becomes possible. The following excerpts from a single interview show this kind of change of personal meanings in progress.

> Mrs. Brown came to the counselor extremely upset over her family relationships and considering the possibility of a divorce from her husband whom she described as a completely intolerable person. After spending some twenty minutes in a veritable tirade against her husband, she made this remark:
>
> "But you know, other people don't see him this way. They think he is a grand fellow and a good father and all that, BUT" and at this point she launched again into a further broadside assault on her husband, his family and everything connected with him. When this had subsided, she said, musingly:
>
> "But he does remember my birthdays and things like that. You know, he has never forgotten once." This was followed by more to the effect that her husband was distinctly not all he might be.
>
> A few moments later, this feeling was expressed with a rather self-conscious giggle: "Golly! I've really been pretty hard on him, haven't I? He really does have *some* good points, you know." A few minutes later, Mrs. Brown says further, "You know, I've lived with that man for ten years now and he still doesn't hang up his pants"—then a long pause as she goes on to say, "But you know, I really think sometimes I kind of like it, looking after him," and she closes the interview with this remarkable statement:
>
> "I guess the real trouble is with me. I'm beginning to think I need this more than he does."

By the use of the technique of recognition and acceptance of feeling the counselor aids his client to clearer and clearer differentiations of the meaning of things for him, while at the same time avoiding any threat to his client. By recognizing and stating these meanings clearly and sharply, he assists his client to further differentiations until, eventually, this process may arrive at those differentiations most troublesome or fear-producing for the client. Sooner or later, the client reaches this point in spite of himself. Once it has been reached, new adjustments become possible.

It may be argued that, if the emphasis of therapy should be upon the meaning of events to the client, then interpretation, questions, directions, and the like should help to speed this process along. It would appear that, if the therapist can perceive these meanings and point them out to the client, the client's differentiations should occur more rapidly. Actually, such techniques may often be disastrous to the client's progress and may impede or even destroy the counseling relationship. Interpretations and information given when the client has not reached a stage of differentiation where he can accept such interpretations are not *his* personal meanings, but his counselor's. As the counselor's interpretations, they may even appear to the client as threatening and force him to defend his position not only against the world at large, but the counselor as well. This failure to accept information into his personal organization is well illustrated by our own life situation wherein we often know what we *ought* to do but don't do it. Until a concept has been accepted into our personal organization, it has little effect on our behavior.

In the permissive atmosphere created by the non-directive counselor and aided by the therapist's techniques of recognition and acceptance of the client's personal meanings, the client is able to differentiate more and more clearly the various aspects of his personal organization. In this process, he is exploring and defining the relationship between himself and the world in which he moves. As counseling progresses, he comes to differentiate more and more clearly the situation in which he is operating, the source of the threats which he has felt, and finally is able to arrive at a new organization of personal meanings resulting in behavior less at odds with the society in which he must operate and infinitely more satisfactory to himself.

A minister's daughter had led an exemplary life for the benefit of her father and his congregation. When she was forced to live her own life in college, away from her family and community, she became deeply confused and upset, and finally came to a college counselor for help. Note, in the following excerpts from counseling, how she struggles with her concept of herself and her relationship to the world about her:

> "A couple of years ago I heard one of Dad's sermons in which he said 'a person has to like himself.' From then on I took it for granted that I did like myself. I decided that I wouldn't change for anything—until this week when I began wondering if I really did. I decided I liked myself but I also despised myself."
>
> "I've decided I'm a two-sided, two-faced person. I've always had to act one way although I felt another. I've always had to be something I'm not. I give appearances but, down under, I'm not that at all."
>
> "I'm not sure what I'm like—I don't know what I am. I'm a man without a country. What is myself? It's funny how sure I was and now I'm not sure at all. I'm afraid of everything at the moment, but I can't find what it is I'm afraid of. I'm afraid to live like this for the rest of my life, but I'm even more scared as to what to do about it. I'm afraid even to think about it. The more I think, the more scared I get."
>
> "My problem is myself. What am I? I'm human, female, five feet seven, period. I want to be sure, but I'm not even sure of myself. Maybe I know what I am but I'm afraid of it. I'm in a panic about myself."
>
> "I never felt I could be myself. I couldn't because of father's job. Now I know I must change myself, but what am I? I must know that."
>
> "*It's a battle between what I think I am and what I really am!*"

Another client, a mother who had been rejecting her children for several years, arrives at this differentiation in counseling:

S.—"I've been struggling with this thing for weeks but I know now what it is. I'm ashamed of myself for not seeing it before and ashamed of myself because that's what it has turned out to be."

C.—"The thing you have been afraid of is pretty clear to you now."

S.—"Yes, it is, but I can't say I'm proud of it. All along I've been upset because of what other people's children were like. My poor kids—when I think of what I've done to them!"

C.—"You feel pretty upset by what other people think."

S.—"That's right. Always it's been like that. I'm afraid I've been pretty selfish in all this."

The Reorganization of Self in Therapy. It is interesting how little the re-evaluation of the self in non-directive therapy has been recognized. Some writers have even gone so far as to suggest that, until non-directive therapy found ways of changing the self concept, it could not truly be considered a fundamental form of therapy. Actually, it is this very redefinition of the self which is the most striking and characteristic aspect of the entire process. Raimy,[3] in his doctoral dissertation, has clearly demonstrated that shifts in the self concept do occur in non-directive therapy.

With a clear differentiation of the relationship between self and the environment about it and with clarification of the nature of the threat to its organization, the stage is set in non-directive counseling for a shift in the self concept. Indeed, such a shift in the concept of self becomes not only possible but almost inevitable. The need of the organism to maintain or enhance its personal organization forces the client to a reorganization of self when it becomes clear that such a reorganization works to the enhancement of the organism involved. The effects of this in non-directive therapy are among the most exciting and fascinating aspects of the process. Since the individual's behavior is a function of the meanings he has given to himself and the world about him, we would expect that changes in the self concept would be accompanied by considerable change in the behavior of the client as well. This is exactly what occurs, in many cases, following non-directive therapy.

One client says near the end of her counseling experience:

S.—"There I was, sitting in the library reading. All of a sudden it hit me. It was the craziest thing—just like that. I thought, 'It's stupid of me to go on like I am. I'm just not a brilliant person and that's all.' It was just all me—mentally! It was so funny. There I was—and I thought, 'Here everybody else is adjusted to themselves and you're trying to adjust to everyone else but yourself.' "

C.—"You feel it's necessary to accept yourself."

S.—"That's right. I know I'll never change from what I am. I was so excited, I couldn't think last night. I thought—*Suppose you had told me what to do?—I'm so glad you didn't!*"

Another client who illustrates this change in behavior was an extremely masculine-appearing young woman who obviously prided herself on this characteristic. Raised in a family in which her brother got all the attention while she was forced to run errands for him, and with no one to play with but boys, the onset of puberty was deeply resented. When an attack of infantile paralysis left her further unable to compete with boys, she set out to prove that she could. She made herself proficient at all kinds of sports, but was never able to achieve an ability equal to a man's. In college, when other girls were looking forward to marriage, she resented her figure and femininity. As she expressed it, "I just can't see the boys I go out with as potential husbands. Pals? Sure. Husbands? No!" She talked in a deep and husky voice, and would rather be seen dead than in anything but slacks and a shirt. In the next-to-last interview held with this young woman, she came to a different organization of her self concept and accepted her role as a woman. She states:

> "You know, I've been trying this weak and helpless angle, and it works. I'm beginning to think I like it. Why shouldn't I? After all—I *am* a woman. I can't be like my brother, so I'll be like me. I'll just have to accept myself as small. I can't do everything—I feel like being smooched with but I don't have anyone to do it with. But that will fix itself. —I'm beginning to feel important!"

Next time the counselor saw this woman he was honestly astounded at the transformation. He hardly recognized his client. She looked feminine from head to toe! Perhaps the best way of expressing this change may be observed in the client's own statement:

> "I'm more feminine now than when I first came in. I feel more feminine and I guess I must act so, because lots of boys that never did before want to help me now. They even want to kiss me and carry my ski poles for me. John especially. I think he likes me better and I know I like him better too—I think I must be on the right track. I feel more comfortable—very comfortable—and—I like myself better too."

The Client-Centered Nature of Non-Directive Therapy. One of the most frequently misunderstood aspects of non-directive therapy is the non-directive therapists' use of the term "client-centered." Some writers[2, 6] have inferred that, since non-directive therapy claimed to be "client-centered," by implication all others were not. This is a most unfortunate interpretation. Certainly, any counselor of whatever philosophy, if he is worthy of being called a counselor at all, is interested in his client's welfare. The term "client-centered" as used in non-directive therapy, is not meant to apply so much to the counselor's concern about his client's welfare, as to the way in which the counselor attempts to see the client's problem *through the client's eyes.* We have repeatedly stressed, in our discussion, the importance of the client's own meanings of events. It is the counselor's concentration on and attempts to understand these personal meanings that is meant by "client-centered." For example, Rogers[5] states, "As time has gone by, we have come to put increasing stress upon the 'client-centeredness' of the relationship, because it is more effective the more completely the counselor concentrates upon trying to understand the client *as the client seems to himself.*" (Italics as in the original.)

This counselor task of seeing the client and his relationship to the world about him as these appear to the client himself is no mean feat. It is one thing to understand this technique; it is a very different matter to put it in practice. The beginning non-directive counselor is very likely to discover, early in his experience, that the seemingly simple non-directive principles become intricate and difficult when the attempt is made to put them to work. He soon discovers that what he knows is of far less importance than what he is. Indeed, much academic knowledge usually required of the psychologist is unnecessary for effective non-directive therapy. This idea has disturbed many critics of non-directive counseling. They have been deeply concerned at what appears, at first glance, to be a renunciation of training. Actually, while it is true that many diagnostic and statistical skills are not essential, effective non-directive therapy requires of the skilful practitioner experience, sensitivity to people, understanding, self-discipline, and a personal growth and development by no means easy to arrive at. While academic training alone may contribute to understanding, it is no guarantee of an effective non-directive therapist. Knowledge of facts can be readily grasped by an intelligent student, but personal growth and self-discipline can only be achieved by most of us through arduous experience.

Recognition of the individual and an absolute respect for his integrity is not just an idea in non-directive therapy. It is a working principle.

Bibliography

1. Combs, A. W. 1946. Basic aspects of non-directive therapy. Am. J. Orthopsychiat. 16: 589–607.
2. Hahn, M. E., & W. E. Kendall. 1947. Some comments in defense of non non-directive counseling. J. Consult. Psychol. 11: 74-81.
3. Raimy, V. R. 1943. The Self-Concept as a Factor in Counseling and Personality Organization. (Doctoral dissertation.) Ohio State University. Columbus.
4. Rogers, C. R. 1942. Counseling and Psychotherapy, Newer Concepts in Practice. Houghton Mifflin. Boston.
5. Rogers, C. R. 1946. Significant aspects of client-centered therapy. Am. Psychologist 1: 415–422.
6. Wrenn, C. G. 1946. Client-centered counseling. Educ. Psychol. Meas. 6: 439–444.

THE THEORY AND PRACTICE OF GROUP PSYCHOTHERAPY

By HELEN E. DURKIN
Scarsdale, New York

JOHN LEVY's Relationship Therapy as we have crystallized, developed, and applied it to groups, is direct interpretative therapy. It is geared to go as deep as but no deeper than the patient's needs and capacities indicate. It is not meant to be palliative, nor to strengthen the patient's repressions.

In the course of a lifetime, people build up various defenses against certain unacceptable instinctual drives, such as hostility. When these defenses are not working well, they become anxious and develop neurotic symptoms (which are danger signals like any symptoms of a physical disease) or they experience general anxiety. This often brings them to the psychotherapist for help. In treatment, they make another attempt to keep away the unbearable drives. Their defenses become *resistances*, and the therapist must dissolve them one by one until the basic drives can be reached, re-evaluated, and handled in a more mature way. This process will strengthen the patient's ego and bring about a real character change.

The Relationship Therapy is a transference therapy, psychoanalytically oriented. Like any such therapy, it is based on the premise that, sooner or later, the patient will bring into his relationship with the therapist the full range of his emotions—his fears, resentments, demandingness, tenderness, and so on. He will express them in the characteristic attitudes and patterns which, under pressure of his early familial experiences, he has built up to defend his ego. For example, as a child, the patient may have been afraid to show his resentment to a tyrannical father. He may have discovered that one way to escape his dilemma and find temporary relief was to run away—physically or symbolically. He will do the same in the treatment situation as soon as resentment is aroused toward the therapist (as a person of authority). Unless the therapist is keenly alert to the first signs of approaching hostile feelings and brings them into the open, he is likely to lose this patient.

Let us examine another similar but more complex example. I was working with a young man who had felt rejected by his mother, toward whom he had a strong attachment. His frustration filled him with rage, but he dared not be angry lest he lose the little he got from her. Both his erotic and his hostile feelings had caused him so much pain that he tried to avoid them. He had started psychotherapy twice and twice he had run away from his woman therapist as soon as he became aware of her as a woman. He left enraged notes behind accusing her of seducing him. It was a kind of triple defense. He projected his erotic feelings on to her, thus getting rid of his anxiety about them. Then he converted his rage at expected frustration into righteous indignation at her bad behavior. Finally, he rid himself of the whole anxiety-producing situation by running away. His people, however, insisted on treatment and he was sent back, this time to me. As I expected, he entered treatment easily, talked with apparent frankness and soon claimed improvement. When he expressed inordinate gratitude, I said, "You know, when you feel so grateful to me, you are apt to become very fond of me, and if this happens you are going to be pretty upset. You may even want to run away." He denied it, but stammered and blushed furiously. I went on, "Patients have all kinds of disturbing feelings toward their therapists. But it is different in treatment from outside, and you will find it is good to talk about these feelings here. This is the way I help people." The boy managed to stay and, instead of running away, he eventually described how he would like to marry me. Again I warned him that, on finding that I have to be impersonal, he would be angry with me and want to leave. Again he was able to stick it out and I was soon listening to a description of how he would like to murder me—with an ice pick. Treatment scenes are seldom as dramatic as this, perhaps never in a group, but the principle is the same.

As you can see, the pivot on which this therapy moves is the analysis of every implication of interpersonal attitudes as they near the threshold of consciousness (whether positive or negative). Trivial as these first evidences of irrational feelings may seem, they are the clues to the patient's ego defenses and, in treatment, these defenses become the resistances. Dissolving the resistance at the earliest possible moment speeds up treatment.

Each patient has any number of such defensive patterns, and, as the therapist exposes one after another of them, the patient gradually comes face to face with the unbearable impulses which were originally behind them. Recognizing them in the controlled interpersonal setting helps him to become aware that they have little bearing on the immediate situation, and to discover their real meanings and sources. Once that has happened, he becomes free to relate himself to people in a more realistic way.

Modification of the feelings of guilt derived from identification with a permissive therapist and the catharsis of dammed-up feeling are a part of therapeusis, but we feel that the greatest change in the patient comes from the gradual dawning of insight within himself as he is helped to recognize how inappropriate his neurotic attitudes are. He catches himself in the act, as it were. The essence of the treatment, therefore, is in the living experience of the relationship.

Since we see our patients only once or twice a week, we do not let them struggle alone with the anxiety they must experience as they become aware of their instinctual drives. We help them to verbalize the anxiety-ridden feelings as these reach the threshold of awareness instead of waiting until the patient is fully conscious of them. This requires skilful timing; for interpreting too soon will *give* him the insight which should *come* from within and will make an intellectual process of treatment. It may succeed only in strengthening his resistance. Waiting too long to interpret, on the other hand, will increase the patient's anxiety between visits when the therapist is not available for help. It would also increase the length of treatment considerably.

These are the principles of therapy, and they remain the same whether used with individuals or with groups, with adults or with children. In the group, the therapist finds himself handling many relationships at once; the feelings of the four or five women (or children) toward one another and toward himself. These start as surface reactions but gradually develop into transferences. The same kinds of defenses and resistances display themselves as occur in individual treatment, and, if one uses a transference therapy like this one, much the same kind of treatment process takes place as in individual therapy.

There are, in addition, many intra-group impacts of which the therapist must be constantly aware and which he must handle. Some women, for instance, express one kind of feeling or another more easily than others and act as catalytic agents for a time. Others will take over, as it were, when other kinds of emotion come up. Although the patient-therapist transference may not become as intense as in individual therapy (sometimes it does), the subjective attitudes of the other patients serve rather quickly to bring into being what we might term secondary transferences.

For example, Mrs. B. may display attitudes that remind Mrs. A. of her mother or, for that matter, of her father. This may arouse tender or resentful feelings toward Mrs. B. which will be expressed in Mrs. A.'s usual patterns. The therapist must interpret such relationships just as he does when they apply to himself, and he must also handle the feelings of the recipient of such transferences. Among such feelings will almost certainly be sibling rivalry, for the group naturally tends to enhance this feeling, which comes out more quickly and realistically than in individual treatment. If the therapist has a tendency to steer away from socially unacceptable feeling, the group will soon develop a sewing circle character. If he sits by, passively but permissively, some superego modification and some catharsis will occur. If he interprets skilfully, timing his comments so that the patients themselves become aware of their conflicts and anxieties, real character change is possible in the group, as in individual treatment. To reach this goal, the therapist encourages all emotional expressions, however socially unacceptable. He helps to bring them out by repeatedly bringing to light the undercurrents of feelings as he senses them behind the factual productions of the patients.

The nature of his comments, as well as his timing, must be geared to the patients' level, psychologically, educationally, and socially, so that they can accept and assimilate what he says. If what he says is too strong for the patients' weak egos (if he deals narcissistic blows), treatment will be vitiated. It follows inevitably, then, that if the therapist's own drives enter into his work beyond the point at which he can recognize and control them, he will create trouble instead of bringing relief. Supervision during training will control this possibility. There is another safeguard, however: such a therapist would soon find himself without any patients—certainly without any groups.

Perhaps the best way of illustrating the method is to try to show it in action by describing a particular group. In order to simplify this too ambitious undertaking, I shall try to follow some of the more illuminating moments in the progress of a particular patient, Mrs. S., within her group, bringing in the others just enough to show the kind of interaction that takes place.

The group I have chosen to discuss consisted of four women who met once weekly. They were mothers whose children were also in treatment at the New Rochelle Child Guidance Center. One patient is an attractive woman of about thirty-four who came because her older child, a girl of seven, suffered from extreme shyness, enuresis, nail-biting, and general irritability. The mother is a neat, modestly dressed woman whose most noticeable characteristic was an almost constant artificial smile which gave her face a mask-like quality.

On the day of our initial meeting, Mrs. S. was the first to arrive and immediately began to complain about her husband and his mother. She spoke stiffly and without feeling, as if reciting a lesson. Her husband, she said, like his mother, never praised but often criticized her, especially for being extravagant. She felt this was unfair because she had always had to be thrifty by necessity. Although her mother and her aunts had been hard up, they were always generous and did not stint with praise. In spite of her righteous indignation, however, she felt that, somehow or other, the whole thing was her fault anyway, perhaps because she was no good as a housekeeper or a mother and did not know how she could ever improve. Since we are, on the whole, passive during the first interview, my only comment was an attempt to help her feel I understood the way she felt. I interpreted not the material, but the undercurrents of feeling, saying only, "I get the impression you have been trying for a long time to keep your chin up." She burst into tears, and spoke with less tension after that.

When the other group members came in, Mrs. S. went right on addressing herself to me alone for a minute. Then she apologized for taking too much time. She was not ready yet for an interpretation of the sibling rivalry she revealed, nor of her need for approval. I waited for an occasion when it was so clear that she herself would see it.

For the rest of the hour the women talked about their children's behavior, as they usually do before they have accepted the idea that they themselves are patients. Mrs. S. listened for the most part, but occasionally she gave an example of similar behavior—however, always choosing it from her own childhood.

Since first interviews are usually revealing as a kind of forecast of things to come, it is a good idea to ask, "What does this all add up to?"

(*1*) Mrs. S.'s artificial, tense manner seemed to indicate that she was struggling to keep an overwhelming anxiety out of sight.

(*2*) What she said seemed to imply, "I am a child, and I'll be good and tell you what you want to know, but please be kind and approve of me."

(*3*) Her disturbance over being criticized made me guess that she used approval and praise to cover up some deep-seated anxiety and warned me that she would want praise from me, too, which would serve to build up her defenses; that she had come for treatment for that very purpose and that when it was not forthcoming she would feel criticized, rejected, angry, and anxious.

(*4*) Her disregard of the other women at the beginning and her quick apology for it revealed her strong drive to keep the therapist (as a mother person) to herself and that one of the first things to handle, as often happens in a group, would be sibling rivalry.

(*5*) Her use of examples from her own childhood confirmed the feeling that she still regarded herself as a child. What had prevented her assuming adult responsibilities would be left to find out. It might also shed light on the deep sense of inadequacy expressed in her inability to cope with household and parental duties.

Since Mrs. S. had not been aggressive enough to keep the floor at the first meeting, I was not surprised when she came late the next time. We find it a useful rule to comment on such evidences of hidden feeling. I said, "I think it was a little harder for you to come this time." She denied it, referring to the silly use of this interpretation in the social-work school she attended for a while. I did not press the point because we feel that arguing with a patient implies self-defense and might make her feel that aggression or resentment is not acceptable. She might, then, shut off future display of it, whereas we want her to feel free to express any feeling whatever. I said only, "You must have been disappointed in the group last time because you had so much to say and so little chance to say it," whereupon she answered, "Well, I did have a lot I wanted to talk about." I knew all the women would be experiencing rivalry at this time, so I said, looking around, "Naturally, every one of you would prefer to be seeing me alone." I could tell by their expressions I had hit home. I said this to relieve them of the need to hide this aggressive feeling, and so to pave the way for later, more explicit expression and interpretation of sibling rivalry. The immediate result in the second interview was that, when Mrs. S. spoke again about her husband, she could include the other women instead of looking at me alone. Months later, when one of the women told how jealous her two boys and husband were of her ministrations, I compared it to their group situation. By this time, they had become aware of their rivalry and could see it in this new sense.

Mrs. S. was early for the third interview and announced that she had made an inexcusable display of herself the week before, when she "blew up" at me. I said that apparently she had felt criticized by me. She said, "I always felt criticized at school when they made that interpretation," and she told us about a teacher who had told a girl to stay away altogether as long as she felt so resistant that she was always late. I said, "I have an idea you are afraid I might send you away too." Mrs. S. nodded, her eyes filling with tears, as she told us how she had always felt on the fringe of her family as a child—lonely and lost. She seemed freer now and dropped the stiff, artificial manner. When she talked about her husband and mother-in-law again, she gave vent to strong resentment.

Apparently, she gained some relief from this tirade, for she soon reported that she was getting along much better with her husband. She had, for the first time, been able to confide in him her feelings about his mother and he had actually taken her part and had been much sweeter to her. She also made him see how she felt about treatment, to which he had been antagonistic. Such immediate results from the first six weeks or so of treatment are not at all rare, for the first layers of anxiety are lifted off. Sometimes the exigencies of the case lead us to terminate treatment at this time, but if the patient needs it, can profit by it, and we have time, we go on.

Mrs. L., a thin, determined-looking woman, who expressed aggression under a usually pleasant, soft-spoken manner, was the woman in the group who had the greatest impact on the treatment of Mrs. S. She was of a lower socio-economic and educational level but had, at the beginning, much more self-confidence. She was typical of many of our group mothers who come only because we insist on it. She saw no relation between her son's anti-social behavior and herself, for had she not always "beaten her brains out" to teach him right from wrong? She said the school principal "had a nerve" blaming her. Very soon she made the demand mothers of this kind always make for advice about handling her son. I said I knew that she and some of the others would probably be disappointed because I do not give advice. Experience had shown that it was best not to if we were to get to the bottom of the trouble. We found that the children's problems were usually related to the parents' emotional difficulties, so that the mothers would talk while I listened and together we would try to find the answers.

Mrs. P., a third member of our group, had so far sat by quietly, saying nothing. She was an untidy, withdrawn woman with a blinking tic. She was a college woman who, because of a deep sense of inadequacy, had married an uneducated man late in life. She was now at a complete loss as to how to cope with her two small boys. She was not really good group material but we had no other time for her and decided to risk it.

Now, for the first time, she took part in the discussion, quoting from books (she was a librarian) to give Mrs. L. the advice I had withheld. I remarked that she seemed to agree with Mrs. L. in that I should give advice. She blushed and said "Well, no, but,—" I smiled in a friendly way and said that here everyone was entitled to her own opinion and that all mothers feel the same way at first. I realized I would have another set of intellectual defenses to cope with here and, furthermore, that Mrs. P. was the kind of patient with so little ego that I must go very slowly in tackling even her resistance. It is best to let such women sit quietly by until, from hearing the others express all sorts of emotions, they have gained sufficient reassurance to be able to touch on their own.

I shall omit the fourth woman from the discussion because she had least impact on Mrs. S. and it is necessary to save space.

Gradually, the women came to talk less and less about their children's behavior. (From then on, when they did revert to it, I would know that some new resistance was making itself felt.) They spoke more about their feelings toward them, their protectiveness, their feeling of responsibility and guilt and, in the end, their resentment and hostility toward them. Some groups are so emotionally mobile that they can even be helped to accept their death wishes in these first months of treatment. This group was far from ready for it. In some groups, where one or two members may go this far while the others would be shocked by it, the therapist must be especially careful to handle the feelings of both kinds of women.

The therapist's part in this group was to help them express these facets of their parent-child relationships freely, to show them the meaning of parental ambivalence and to help them see that they were not alone in their various predicaments.

During this period, Mrs. S. was the chief catalytic agent, leading the other women gradually to more personal problems because of her emphasis on her own childhood. Mrs. L. responded by telling of her childhood experiences, too. She criticized her mother freely. In her characteristic way, she showed her sibling rivalry by aggressively taking the floor most of the hour. She complained of being snubbed and ignored by so-called friends for whom she tried to do too much. Gradually, she began to see that she was not just the victim of their hostility but that she herself activated it by the chip on her shoulder. The same situation appeared to hold in her relationship with her son and, as a result, she began to treat him more tolerantly.

Mrs. P. talked freely at this time, too, for she half lived in her childhood fantasies anyway. Becoming aware of the meaning of some of them (her driving competition with her mother) seemed to bring her into somewhat closer touch with her daily life.

Since the others gave Mrs. S. little chance to talk these days, another of her defensive patterns was high-lighted. She appeared to be listening sympathetically. When she finally got the floor she complained of one of her severe headaches. She thought it was because someone had called her as she was hurrying to get to our meeting. It was the head of a committee who told her a long story of woe in order to convince her that she should take over some onerous duties. Mrs. S. had listened with apparent sympathy, although she was angry. She took the job, but now felt overburdened and inadequate. I said I had an idea she was feeling the same way about the group lately when she sat for such long stretches listening with apparent interest to the other women's troubles. I knew this must be frustrating to her. Mrs. S. laughed a little and said she had also been having headaches after our meetings lately. I tried to get her to talk about these headaches, from which she had been suffering all her life. She started talking about them, but soon drifted off into a long story about how angry she became when her husband insisted on taking the whole family on endless drives and then expected dinner in a jiffy after she got home. As a matter of fact, she said, she often had headaches on Sunday. This demonstrated the feeling behind the headaches, and she laughed again as she put two and two together. Then she suddenly recalled an illuminating childhood memory. When she was about nine years old, she used to go to an art class. The other children used to make many demands on the teacher whereas she sat in a corner not daring to ask for help, although she could never draw as well as they. When she went home, she used to cry herself to sleep. I asked why she thought she had just recalled this. It was easy for her to see that she was doing something very similar here.

From that point on, Mrs. S. became a little more aggressive in the group. Soon there were several signs that her hostility toward her mother, of whom she had said only the most complimentary things so far, was mounting. She again became anxious and resistant, even though she had been somewhat prepared for the recognition of such feelings in herself by Mrs. L.'s open, bitter criticism of her mother.

When her resistance had been broken through, Mrs. S. revealed a negative transference to Mrs. L., whose overprotection of her son reminded her of her own mother's attitude. At this time, the group was talking about how to tell children the facts of life. Mrs. L. revealed that she could not talk about it to her son; that she worried about whether he masturbated; sometimes she even spied on him. Mrs. S. looking angry and upset, I said, "You seem to have some feelings about this, Mrs. S." In a cold, intellectual way, she said that a child's reaction to such handling could only be one of rebellion. I said, "I think you are feeling irritated at Mrs. L. for the way she treats Charles." She broke through then and said angrily, "Yes, he'll never forgive her, you know," to which I replied, "This seems to be the way you feel too; I think she reminds you of your mother." She burst out, "That's right, I'll never forgive my mother. She always pretended that sex was all romance and then one day she called us children together and told us the 'bestial side of sex' all at once. She made it sound so gruesome I never have got over it. It just about spoiled the first part of my marriage"—and as usual she cried.

From then on, her mother came in for one bout of criticism after another, until she reached a climax when she was telling us about her mother's irritating "sweetness and light" attitude when she came to take care of her grandchildren. Grandma always sidetracked them when they were going to be aggressive, and Mrs. S. felt that was at the bottom of her little girl's shyness and inability to be aggressive, just as it was at the bottom of her own inability to assert herself. Her voice rose in a crescendo as she burst out. "Pretending, pretending, that's all my life has ever amounted to—no wonder I can't manage my household, my children, or anything." She began to cry bitterly. While she was working through this hostility, her real relationship with her mother, who was visiting her, was very much disturbed. She had things out with her mother regularly. Her husband called me to ask what was going on, and I assured him that her reactions were an unavoidable part of treatment, but I felt they would soon subside and she would find a better equilibrium as a result. Fortunately, it was not long before Mrs. S. reported that she was much better able to manage her children now and that her daughter was much more aggressive than she ever had been. The incidents she told seemed to show that Mrs. S. was asserting herself in a more adult way at home. The explanation seemed to me to be that as soon as she was able to face aggressive feelings toward her mother, the need to remain a child in order to suppress her feelings of hostile rivalry toward her was greatly reduced. She could be the mother to her own children now instead of being just another guilty child with them. She was also getting on better with her mother-in-law, who no longer was the recipient of hostility meant for her mother. Her husband also apparently responded to her increased maturity, for she told us he had been so attentive she felt as if she were a bride again.

By that time, most of the women had talked a lot about their resentment toward the various members of their family, but they had not yet been able to criticize me directly. As in all matters of aggression, Mrs. L. was the leader. One morning, when they were all resistant at once, the conversation turned toward teachers. Mrs. L. (to whom I had recently had to deny information about what her son had told his therapist) said, "Teachers never tell you anything. When I want to know something, I ignore them and go straight to the principal," and Mrs. S. chipped in, "I think they ought to learn some psychology." (Note the appeal for my approval.) "They think they can handle all children the same way" (objection to group). Mrs. P. complained that, at school meetings, the teacher never seemed to notice her, so that she never had a chance to talk to her about her little boy.

Enough material about their familial backgrounds had come out by this time so that I could easily show them all that this was exactly how they had felt about their mothers, persons in authority, and, more recently, about me. Each one, in her characteristic pattern, had presented a picture of her own irrational idea of my attitude toward her. I showed them each the meaning of what they had been saying. Insight was stimulated because each woman could see so clearly that what I said about the others was true, and they all laughed. Because the whole group was expressing resentment toward me at once here, it was possible for each to accept such feelings in herself at this time (security from the gang).

Following this episode, Mrs. S. expressed resentment to me more directly, but the persistent struggle of her need for approval over her emerging hostile feelings made treatment a slow, laborious process for her. The insight she gained was partial and frequently repressed again, as often happens. One episode threw into clear relief the way she used an appeal for approval as a means of defense against anxiety about hostile feelings. She was expressing resistance again one day, by talking about her dislike of doctors, when she stopped abruptly and asked me if fear could be repressed. I said, "Suppose you tell me what you mean." She told us about an operation she had undergone several years before. She had been calm beforehand and was often complimented by the day nurses. The night before the operation, however, she could not go to sleep, and one of the night nurses scolded her severely for it. Suddenly she became panicky and the next morning she insisted on calling up her doctor to ask him some questions. He told her abruptly that she was afraid. She referred to her calmness until this time. He answered harshly that she had only repressed her fears. She was angry, but helpless and more panicky than ever about the operation. She felt he had had no business making interpretations. As she talked about it I helped her see that she had been all right until someone had criticized her, as if being approved of (loved) had been a way of keeping down panic. I showed her, too, how the same pattern was working in the group. She had often tried to win my approval (this much had come out frequently and she had long since accepted it), and when I did not give it to her she became angry, panicky, and resistant.

It was clear now why Mrs. S. had come to treatment. Ostensibly, it was to learn how to handle her affairs. What she wanted, unconsciously, was to get approval from an authority so that she could continue to suppress hostility and deeper fears. It is clear, too, that had the therapist given her advice or praise he would have strengthened her defenses and made it harder, if not impossible, to work through the anxiety-laden drives that lay behind them.

When she was a child, her family's defensive generosity with money and praise had served to keep up her defenses, although they often wore thin and she had actually had a nervous breakdown at college. When she married, her husband and critical mother-in-law withheld both money and approval, and the defenses broke down. She was both angry and filled with a deep sense of inadequacy.*

When the same pattern had played itself out in treatment, the vicious circle could be broken because the therapist neither played into her defensive patterns nor rejected her, but showed her the meaning of what she was feeling.

---

* Her basic fears had to do with castration anxiety which had frequently been indicated, but which she had certainly not been ready to touch so far.

What had kept her in treatment so long in spite of the enormous struggle she had to undergo to face her hostility, was a deeply passive dependent drive. I was her last hope of help and she had to stick to it. (Had the therapist been punishing or masochistic, she could not have stayed.) Her passive trends had also been evident from the very beginning and, by this time, she had gained some understanding of them. But she could not experience the tender and erotic nature of the feelings connected with the dependent drive until she was released by the expressions of resentment in a situation which did not bring with it the rejection or the retaliation she feared. I will not attempt to describe the slow steps by which she became aware of what she called "wanting to lean." The incident in which she recognized the connection between this drive and her hostile feelings, however, is worth recording because, without it, it is not possible to understand how she came at last to grapple with still deeper layers of anxiety.

Toward the end of our second year of work (about the 60th interview), when the others were complaining about their children again, Mrs. S. joined them, this time without any attempt to avoid my disapproval (heretofore, she had always worked in some compliment to me, for safety). She implied that her little girl was worse now than when she had first brought her to the center (in spite of the fact that she had already told us that the nail-biting and enuresis had stopped). I was glad to see that hostility to the therapist was stronger now than her need to side with the mother (myself) against the siblings as she used to do. Here we see a decrease in dependency.

She denied my interpretation of her doubts about treatment, but the next week she came in feeling depressed. She was worried, she said, lest she and her husband would have to help support her aunts who were now trying to live on their pensions. Using a deadly monotone, she gave the minutest details about how these aunts had helped her out financially, through college. Now she felt mean to begrudge them help.

I had only to remind her that this tone of voice usually meant resistance, and she immediately spoke in a more heartfelt way. The youngest of these aunts was very important to her because she had been the only bright spot in Mrs. S.'s life the year her sister had been sick. The rest of the family's attention had been concentrated entirely on the invalid, and only this aunt had paid any attention to Mrs. S. It was possible to show her how she was struggling here with guilty feelings about being hostile to someone she wanted to lean on. I showed her, too, how she was feeling just that way toward me at this time—needing my help and feeling guilty when she was angry or resistant so that she did not want to talk in the group. She indicated her growing insight by giving examples of how this same interplay of feelings had often happened in her relationship with her mother, although she had never before known what it was all about.

The following week, Mrs. S. came in a warm friendly mood such as she had never shown before. She thanked me and said she was touched by my desire to help her. She spoke tenderly of her aunt who, she said, reminded her of her father: they both loved church music. Even now she gets a thrill whenever she hears organ music. Soon this was her theme for the hour. It was the first time she had said much about her relationship with her father, and the first time that genuine tender, erotically tinged feeling came out. Aunt, father, therapist seemed to be equated here.

Apparently, it had been even harder for Mrs. S. to express affection, which was choked off by such angry feelings that she feared rebuff and even retaliation. The praise her mother had offered her had merely covered up basic rejection and preference for Mrs. S.'s brother. Our patient had, therefore, found herself in constant need for an expression of love, or its substitute—approval. Never satisfied, she kept asking for more, repressing the hostility which she felt might cut her off altogether. Now, at last, she could express love and she looked much happier as a result.

At the last visit before vacation, Mrs. S. expressed regret at leaving, but compared herself favorably to the year before, when she had been terrified of the summer vacation. She had an excellent summer and when she returned in the fall she seemed an entirely different patient, for she plunged with real feeling and spontaneity into some of her underlying conflicts. It was apparent that we were through the worst resistance. She tackled first her difficult, partially frigid sexual relationship with her husband, which heretofore she had declared was one thing with which she needed no help. Next, she came face to face with her inability to assume either a feminine or a masculine role (which was the real significance of her intellectual strivings). Trying to compete with men made her feel hypocritical and "hollow." She always felt she might be found out at any moment. On the other hand, being a woman always made her feel dirty and castrated. Working out some of the hidden feelings and fantasies surrounding this basic conflict gave her enormous relief, making it possible for her to express franker sexual feelings toward her husband. She has been looking radiant of late and speaks of experiencing a new "*joie de vivre*" entirely unknown to her. Her housekeeping is improving at last, and the family manages to have more fun when they are together. It seems clear that, by the end of the year, she will have worked out a much better relationship with her husband, for he seems to be responding to her own changed attitudes.

In spite of the fact that I have selected only a few of the highlights of a long, tedious process, it seems clear that treatment is not a series of beautiful insights. It is more like putting together the tiny parts of a jig-saw puzzle. The fact that there are three other women to be dealt with makes the puzzle more complex and one is frequently distracted from one by the others, but, on the other hand, they offer many clues and stimulation that would never occur in a two-way relationship. This therapy offers even the beginner a safe guide through what often seems a dense fog. If he concentrates on sensing and handling only the immediately underlying feelings that the patients are experiencing at any given time, he will succeed in gradually building up a clear picture and helping his patients to richer personalities and fuller lives.

PSYCHODRAMA AND GROUP PSYCHOTHERAPY

By J. L. MORENO
Psychodramatic Institute, New York, N. Y.

IN the last two centuries, three revolutions have taken place in the field of psychotherapy. Each was characterized by a specific change of operation. In each case, the new practice gradually compelled an overhauling of theory, but in each successive case the new method was broader in scope and, to a degree, included the previous form.

The first, in the middle of the eighteenth century, is connected with the name of a Viennese physician, Mesmer. The operation, *hypnosis*, consisted in putting the patient into a state of trance. Mesmer thought that the hypnotist is responsible for the state of the "hypnotizand" and developed a theory about animal magnetism according to which a fluid travels from the physician to the subject.

At the end of the nineteenth century, another Viennese, Freud, brought about a new revolution by discarding the hypnotic sleep as a means of treatment and establishing another form of operation. Patient and physician faced each other in full consciousness, the patient was told to tell the doctor whatever came to his mind. The physician expected to attain by this method, which he called *psychoanalysis*, all the results which had been attained previously by means of hypnosis and many more things to which the hypnotized state of the patient closed the doors. The psychoanalytic method of operation brought Mesmer's theory of animal magnetism and all its intellectual modifications *via* Charcot, Bernheim, and others into discard, and it was replaced by the well-known system of psychoanalytic theories.

During the crucial years between 1900 and 1925 in which psychoanalytic theory and practice developed, there have been many widely discussed differences between psychoanalytic schools. However, the conflict between Freud, Jung and Adler was due to different views of analysis and interpretation, and there was no conflict between them as to *operation*. The Freudians emphasized libido and its cathexis as chief determinant of human behavior, while Adler preferred inferiority organs and inferiority feelings as the core of his analysis, and for Jung it was the collective unconscious and the extrovert-introvert types of personality which seemed to matter. But if we could have entered the office of a Freudian, an Adlerian or a Jungian between 1910–1930, the operation would have been about the same: a physician and a patient alone, in a doctor's office strictly private and sealed from observers. There were slight modifications: in one case a patient relaxing on a couch, in another case facing him, the patient sitting in front of the doctor, in still another case the procedure being more informal and the duration of the treatment shorter. But in all cases the patient would have been found talking freely about himself and the physician giving an analysis of the material elicited. However great the contrast may have been in the ways of interpretation and in its depths, there was no difference in operation.

In our own time, in the last twenty-five years, a new revolution took place when the first therapeutic theatre was started in Vienna. It was again due to a radical change of operation. The method has become known as psychodrama, sociodrama, role playing, and action taking. *The patient is now an actor on the stage, acting before a smaller or larger audience of other patients.* The physician-patient relation has become subsidiary. Again, we are in the midst of an overhauling of theory. With the new operation, new concepts and theories are emerging. It consists of two procedures: (*a*) treatment of the audience (group psychotherapy); (*b*) representatives of the group portray on the stage the problem from which the audience ails (action therapy). The group is facing the mirror of itself (in many versions) on the stage. It looks into this mirror and sees itself. The responses coming from the shock to the audio egos (members of the audience) and to the auxiliary egos (actors on the stage) are systematically followed up.

DISCUSSION ON GROUP THERAPY

DR. S. R. SLAVSON (*New York, N. Y.*):

DR. DURKIN's paper* demonstrates the very significant fact that group psychotherapy is not different in any essential respects from individual treatment. We have seen in her paper how the focus of therapy is the individual patient and that the group is used only as a tool in treatment, rather than being its center.

Group psychotherapy must be considered as only one form of psychological therapy and is an integral part of it. It bases itself upon the same understandings and employs precisely the same dynamics as does individual therapy. However, the presence of more than one patient and the interstimulation that occurs as a result of this, aid or retard the therapeutic process for each in accordance with the situation. As individual psychotherapy, group psychotherapy is also based upon five dynamics: (*1*) relationship, (*2*) catharsis, (*3*) insight or ego-strengthening, (*4*) reality testing, and (*5*) sublimation.

Relationship and transference are too well accepted and understood to need elaboration here. Catharsis as it appears in group therapy, however, is somewhat different. Verbal catharsis is employed in interview group therapy, while activity catharsis prevails in activity groups. Children of pre-school and school age can express their feelings and attitudes through action more freely and appropriately than through discussion. While insight in group psychotherapy is not as profound and deep as in psychoanalysis, the aim is to help each patient in the group to gain some understanding of his motivations and behavior in terms of unconscious conflicts and strivings, as well as in the light of his emotional history. This latter fact has been particularly well demonstrated in Dr. Durkin's paper.

In our own work with activity groups, we found that even in such groups where no discussions are held or interpretations given, children become aware of their own changes in attitude and alterations in feeling tones. Frequently they verbalize this, although most often it remains un formulated in words. This understanding is evidenced in many ways even when it is not verbalized, and has been described as *derivative insight.*

Reality testing is the fourth dynamic operative in all psychotherapy. In fact, in some respects, this is the greatest contribution of group therapy to the therapeutic process. Each individual tests himself against reality in everyday pursuits and contacts, and patients meet the pressures of the milieu in various ways. They observe how adequately they deal with reality and how much they are hurt by it. In accordance with his basic problem, the patient may become latently hostile, overtly aggressive, or he may withdraw. As he finds himself adequate or wanting, he returns to the therapeutic situation to examine himself or gain support from the therapist. To give this support is an important function of the therapist, whether in individual or group treatment.

---

* See pages 889–901 of this monograph.

The group is a more tangible reality than is the individual interview. Here are persons who also have problems, and the best results are obtained where the psychologic syndrome (though not necessarily the symptom) is similar in all the members of the group. There are present hostilities, rivalries, jealousies, antagonistic and ego cravings, mutualistic support. The patient finds himself in a realistic situation and can test himself within the therapeutic situation itself. Thus, therapy and reality are fused into one. Although this unitary relation is not essential to obtain results in therapy, it accelerates the process greatly. There are patients to whom this confluence of therapy and reality is helpful, if not essential. The very important point to be borne in mind, however, is that group therapy is in essential respects the same as individual psychotherapy and that it is based upon the same concepts and dynamics.

In group therapy, transference (relationship) is greatly modified through the network of emotions in the group. In addition to the attitudes toward the therapist, there also exist sibling relations and various identifications. Thus, the transference is *diluted*. Catharsis emanates not only from the transference relation but it is also stimulated by the other patients, and by the fact that anxiety is less intense in a group.

One point needs to be stressed and this is the factor of *ego-strengthening*. While ego-strengthening does occur in all therapy, it is the very foundation of activity groups. The child brings his impulses under control through the pressures of the group and through the fact that he gains status and acceptance. To the child, whose character is not fully formed and set, experience takes the place of insight. We found that insight for the young child is not as important as are release and strengthening of his ego.

Another dynamic that helps the therapeutic process in a group is *target multiplicity*. The therapist is not the only one who receives the hostility and love of the patients. Feelings are displaced on or redirected to other members of the group. In activity groups, the child destroys objects in the room or he attacks other children to displace hostility toward the therapist. In interview groups, destructive attitudes and emotions are redirected toward other members of the group.*

Another important dynamic that is present in therapy groups as described by Dr. Durkin can be termed as *cathexis displacement*. This is the dynamic in which emotional ties are established with the therapist, other members of the group, or the group as a whole. The outcome of this process is that the patient becomes more emotionally free as his earlier ties are disengaged from their infantile anchorings.

Since each patient feels that his problem is not peculiar to himself and that others are in the same situation as he, his deflated self-evaluation is repaired and his ego strengthened as a result. He is thus able to deal with himself and the world around him more effectively. This we have described as the process of *universalization.*

---

* See SLAVSON, S. R. Differential dynamics of interview and activity group therapy. Am. J. Orthopsychiat. April, 1947.

Dr. Moreno's contribution to the field of psychodrama need not be reemphasized here.* Its value for specific patients has been amply demonstrated and many of his formulations have now been incorporated in the general psychiatric literature. It will be of great interest to examine further the nature of the relationships in the psychodramatic situation in terms of group dynamics, which I hope will be undertaken in the near future. This should prove a very fruitful field.

---

* *Cf.* pages 902–903 of this monograph.

WHAT DID THE CLINICAL PSYCHOLOGIST LEARN FROM THE WAR?

By MAX L. HUTT
*Teachers College, Columbia University, New York, N. Y.**

IF we attempt to answer the question contained in the title in terms of the evolution of entirely new techniques of measurement or in terms of startlingly new concepts of personality or mental disease, we shall find that our search for such developments is a vain one, for, by and large, so far as the writer knows, no great discoveries in clinical psychology were uncovered during the war. Nevertheless, the many psychologists from all of the branches of military service with whom the writer has talked, and the many clinical psychologists with whom he was privileged to work in the army, would all agree that a great deal was learned. Nor would they be content with an assertion that the result of this learning process was indefinable, for there were many specific and concrete gains. The difficulty in specifying these gains is that some of them were gains in a negative sense, and another is that the gains were not always in terms of tests and devices but rather in techniques and approaches. In fact, one can sum up our war experience with two complementary, although *apparently* contradictory statements. The first is the humbling conclusion that the scope of our ignorance in the field of clinical psychology is still appallingly vast. The other is that, as the result of the pressing needs of the military situation and of an unprecedented array of clinical data, many important, if not epochal, advances were made in technique and some in theory.

If our previous training and experience did not yield completely satisfactory answers to the needs of the war situation, they did at least enable us to get adapted to these needs in many ingenious and practical ways. We shall attempt to discuss these advances made during the war under four major headings.

Extension of Basic Clinical Experience. The experience of psychologists who worked in the Medical Corps, and especially in the Neuropsychiatric Division, was broadened and deepened in at least three ways. The most obvious of these was in the extension of clinical practice to a large number and a great variety of neuropsychiatric and special medical cases. These cases differed from those usually seen in state institutions or in community clinics. They were essentially non-institutional cases, young adults representing a cross-section of the healthiest segment of our population. There were relatively few chronic neurotics and deteriorated psychotics but there were many acute neuroses and neurotic reactions and many acute or incipient psychoses. In civilian life, we rarely had the opportunity of examining or treating such people. They were in the borderland of psychiatry or even beyond these borders and rarely came to the clinic or hospital. In addition, however, there were many supposedly non-neuropsychiatric cases, the many general medical cases who manifested psychosomatic disturbances or were "cured" medically but did not get well. There were also hundreds of cases of diagnosed aphasia, hundreds of paraplegics and thousands of amputees, blinded, deafened and the like, as well as large numbers of cases with recent brain damage.

---

* **Present address: Department of Psychology, University of Michigan, Ann Arbor, Michigan.**

Clinical psychologists had worked with few such cases, from the young adult population, prior to the war, and they soon learned that the previously used tests, differential signs and clinical symptoms could not be applied directly to these individuals. Moreover, they were confronted with individuals coming from all parts of the country, with extreme variations in education, social experience, and occupational background. These background variations may have been known to psychologists from their clinical literature, but few of them had ever been confronted by such variations in their home communities.

A second aspect of the extension of clinical experience was in the field of guidance and therapy. Although in many cases psychologists were unprepared or only partially prepared to offer therapy, the tremendous number of cases requiring therapeutic care compelled the use of clinical psychologists in therapy. In a survey made by the writer during the last year of the war, it was found that clinical psychologists assigned to army medical installations were devoting, on the average, 25 per cent of their total duty hours to psychotherapy. Their roles in therapy varied largely in accordance with their training and interests and, to a considerable extent, with the attitudes of the psychiatrists to whom they were responsible. Most therapeutic work was on an individual short-term basis, averaging 5 to 10 sessions per patient. The second largest type of therapeutic aid given was in group therapy, which was explored more fully in the war than had ever been the case previously. Some psychologists directed comprehensive rehabilitation programs involving the use of fairly large numbers of ancillary workers.

Such diversified therapeutic programs offered psychologists opportunities they were often denied in civilian practice. They had to learn while they worked, and learn they did in more ways than one. It is not possible to discuss in this report the varied types of therapy employed, because of lack of space. Suffice it to say that, despite the extensive experience many psychologists had had in the Army, most felt that they could profit from further systematic study of this field, supplemented by appropriate extensive clinical experience. As an aside, it may be indicated that group therapy, and by this I do not mean simply group mental hygiene discussions, was found to be a much more worth-while procedure than many had anticipated, and in some cases group therapy became the therapy of choice.

A third type of extension of clinical experience was found in working as a member of a neuropsychiatric team. To many psychologists, the team concept was a new one. By army directive, psychologists became members of a team consisting of at least a psychiatrist, a psychologist, and a psychiatric social worker. These members had to learn to work together, to supplement each other, and to accept designated responsibilities.

There was, and still is, considerable difference of opinion, and even confusion, about the team concept. In some cases, personal and professional jealousies jeopardized the functioning of the team and did not bode well for the patient. On the whole, however, the concept of the neuropsychiatric team, which did not arise in the war but was certainly strengthened during it, was accepted, and the teams functioned harmoniously. Partly as a result of the war experience, the main problem today seems not to be whether to accept the team concept in neuropsychiatric practice, but rather defining and analyzing the concept more adequately.

The extensions of clinical experience which we have discussed suggest important implications for training and research, which are not, however, within the scope of this paper.

Psychometric Problems. No attention will be given in this report to the numerous screening tests and the special trade and aptitude tests and other assessment procedures developed by both the Army and Navy. Noteworthy as these methods were with respect to theory of testing and methods of test construction, they were not essentially clinical instruments. We shall confine our inspection to individual clinical devices.

In the Army, according to a survey made by the writer, clinical psychologists devoted about 30–35 per cent of their time to individual testing of intelligence, personality, special disabilities, and the like. This was their most important function in terms of both time and significance. However, the psychologist was no longer a mere psychometrist. He was essentially a psychodiagnostician. Numerical scores on tests assumed their rightful place in clinical work as minor aspects of the results of such testing. On the whole, the very marked emphasis was upon understanding the psychodynamics of the patient. To this end, both objective and projective tests contributed significantly. In neither case, however, was the calculation of a score, the construction of a scattergram, or the summation of the quantitative aspects of the test protocol an end in itself. These became the bases upon which cues about the personality or hypotheses about the patient's difficulty could be developed. Some cases were referred, of course, for determination of possible mental deficiency, where some quantitative results were particularly important. Even here, however, much more was usually demanded.

Psychologists made increasing use of observations of the subject during the test situation and made qualitative analyses of test responses.

In a sense, this was forced upon the psychologist, not only by the insistent requests for differential diagnosis, dynamisms of adjustment, and leads for therapy, but equally as much by the plain fact that the standard test profiles, scattergrams and signs *did not fit* the cases being referred for evaluation. The nature of the problems encountered and the cases referred has already been noted in the previous section of this report. These problems were unique. Old normative data were often inadequate. Hence, the more expert psychologists relied more and more upon truly clinical evaluation of test and observational data. When this was not done, serious errors in evaluation often resulted.

Hence, it followed that intelligence tests, while used very frequently, were more often analyzed for evaluation of the personality than for a rating of intelligence. The Army Wechsler Test and later the Wechsler-Bellevue Scale was the test of choice because it lent itself to such evaluations. As already noted, scatter analysis went an important step beyond simple test ratings, but it did not go far enough. The intelligence test was analyzed for content of responses and for clinical behavior elicited. Dr. Hunt has stressed the importance of this type of evidence in his recent paper in the *Journal of Clinical Psychology.** The writer wishes to confirm Hunt's conclusion that there is far too little published material concerning the qualitative data yielded by the individual intelligence test. This lack becomes increasingly important in view of the developing function of the neuropsychiatric team and the increasing emphasis placed upon the diagnostic skill of the psychologist.

Projective tests of many kinds were used extensively during the war and were relied upon more heavily than any other type of test for intensive personality evaluation. It is the writer's opinion that this type of test will be used much more extensively as a result of the war experience and that a great deal of research with many kinds of projective devices will be done in the near future. The possibilities of analyzing not only personality structure, but also specific complexes, conflicts, ideations and attitudes are indeed tremendous. We are only on the threshold regarding our use of this method. While much of the data yielded by such tests can be obtained in intensive psychotherapeutic work, these data can often be secured more quickly and more completely by projective analysis.

With respect to the problems of aphasia and mental deterioration, our military experience has convinced many that the psychological tests for these conditions are hardly adequate. On the basis of this experience, several studies have already been undertaken to gain a better understanding and to develop more suitable tests than were heretofore available.

Clinical Procedures. We have already alluded to changes in clinical procedures in previous sections of this report. Only one aspect of this problem will be discussed here. There appears to be a fairly sharp division of opinion concerning the organization of test schedules or batteries of tests to meet the needs of clinical work. One group suggests the routine use of a comprehensive battery of tests for all or most patients. This battery is supplemented occasionally by special tests in exceptional cases. Some of the arguments in favor of such a procedure are: (*1*) comparable data on a number of tests become available for research; (*2*) a routine battery of tests offers a systematic check on the most important phases of the patient's personality; and (*3*) the psychologist gains familiarity with the battery and is more readily able to integrate the findings successfully. These advantages are real and important. However, the war experience has convinced the writer that such a procedure generally is wasteful of both the patient's and the examiner's time and that it does not provide for the more detailed probing required in psychiatric illness. This experience suggests the advisability of tailoring the battery to the individual requirements of the case. Different kinds of cases require markedly different batteries of tests. It is a gross waste of time to test routinely with the entire battery for all cases. Moreover, special attention can be given to those aspects of the problem needing further analysis. In some cases, indeed, only one or even no tests at all need be given. If equal amounts of time are expended with each method, the latter approach, requiring careful prior analysis of the needs of the particular patient, will probably reveal more important data for the time allotted, since only the area about which some question is raised will be explored. There is also the danger in the routine use of a battery that the approach will become mechanical and significant clues neglected. The flexible battery invites the development of sensitivity to the needs of the case and of adaptations of tests and test batteries to fit those needs. In any case, whether one accepts the first or the second view, it is clear from our experience during the war that we can expect the development of new types of test batteries for varied clinical use.

---

* See HUNT, W. A. 1946. The future of diagnostic testing in clinical psychology. J. Clin. Psychol. 2: 311: 317.

Clinical Theory and Nosology. Relatively little that is new was learned concerning the theory of psychiatric illness. In the main, the contribution of the war experience was to highlight certain conceptions and to afford an opportunity of testing some therapeutic approaches, *i.e.*, shock therapy, narcosynthesis, hypnoanalysis, and the like. Although they had been known previously, the war sharpened some concepts considerably. For example, the distinction between "traumatic reactions" and chronic neuroses became much clearer and the distinctions between "acute psychotic episodes" and "essential psychoses" became more evident. The effects of recent traumatic brain damage were studied intensively and on a wide scale. The psychologist contributed to the charting of these distinctions, and much of the exploratory work remains to be published. It was as if a great experiment in evaluating the effects of sudden and of prolonged trauma upon personality structure and function was being conducted. Conventional psychiatric labels were often wholly inappropriate, so that new nosological classifications had to be and were developed. The role of certain etiological factors in breakdown, especially the breakdown of the so-called "normal adult," was studied, if not systematically, at least extensively. As classification of mental illness improved, diagnostic techniques also tended to improve. Some of these advances have already been reported fully in psychological and psychiatric literature in the past 2–3 years and need not be repeated here.

Summary

The impact of the war upon clinical psychology has been great, although no startlingly new developments took place. The extension of the boundaries of clinical psychology, the deepening of experience with tests as diagnostic instruments and with psychotherapy, the adaptation of clinical techniques to meet diverse problems of new types, and the improvement in our knowledge concerning the theory of mental illness are all resultants of this impact. Inadequacies in our present clinical knowledge and, in particular, in our training programs have been emphasized. One thing more needs to be added: The profession of clinical psychology, as a professional group, was ill prepared to meet the needs and the opportunities created by the war. Clinical psychology was not represented at the highest levels of military command until late in the war, and even then not adequately. However, the war has served as a great catalytic agent in the development of clinical psychology, and its effects will be evident for many years to come.

FUTURE TRAINING IN CLINICAL PSYCHOLOGY

By JAMES G. MILLER
Clinical Psychology Section, Veterans Administration, Washington, D. C.

THE tremendous increase in demand for the services of professional personnel qualified to take part in the care of neuropsychiatric illnesses results from an increased ability, on the part of both physicians and laity, to recognize the presence of these disabilities; from a greater willingness of the general public to admit that they are suffering from such disturbances—a consequence of the general educational campaign which has caused them to be viewed more understandingly; and also perhaps from changes in the cultural patterns of this modern age which may serve to raise the incidence of such illnesses. Whatever the causes for this demand may be, it is clear that the medical profession cannot adequately meet it alone. The services of clinical psychologists, psychiatric social workers, and related professions will be needed in very large numbers, besides the skills of neuropsychiatrists.

Before the war, clinical psychologists were employed sporadically in private hospitals, clinics, and guidance centers, as well as in state hospitals and clinics. Only a handful were employed by the federal government. In private and state institutions, the demand for clinical psychologists is growing rapidly, but the most dramatic outgrowth of the war has been the expansion of large government programs in this field, particularly in the United States Public Health Service and the Veterans Administration. Eventually, the program of the Public Health Service in its central and decentralized aspects may well be larger than that of the Veterans Administration, since it has responsibility for the mental health of seven-eighths of our population. The Veterans Administration has responsibility for only one-eighth, but its rapid expansion immediately at the end of the war has enabled it to get under way earlier.

It is expected that clinical psychologists in the Veterans Administration will be employed in at least five sorts of installations—general medical and surgical hospitals, neuropsychiatric hospitals, mental hygiene clinics, aphasia centers, and paraplegia centers—the latter two located in general medical and surgical hospitals. In these installations, they will work in the neuropsychiatric team, sharing a large part of the responsibility for diagnosis, research and therapy in connection with the psychiatric and psychological care of the veteran. Intensive efforts have been made by the Veterans Administration to recruit clinical psychologists, but these were doomed to poor success since the organization could effectively employ at least twice as many qualified clinical psychologists as there are in the whole country. The only solution to the great shortage in this field was obviously the establishment of a training program.

The Veterans Administration has therefore determined upon the policy of sponsoring a large-scale training program leading to the doctoral degree in clinical psychology in collaboration with those universities accredited by the American Psychological Association in this field. In 1946–47, 22 such universities cooperated with the Veterans Administration, taking 218 trainees. In 1947–48 the number of universities is 36 and the number of trainees 468.

Duties for which the Neuropsychiatric Team must be Trained. The training in psychological fields made necessary by the clinical programs now developing may reasonably be discussed at two levels: first, what would be desired if circumstances were entirely favorable; and second, what can be accomplished out of the realities which prevail at present. In all likelihood, it will be possible gradually to improve the existing situation so that it approaches the ideal more nearly.

Let us first consider what would be the most satisfactory of all possible training centers in the field of professional psychological studies. At first, it would seem to be where every student receives full training in at least psychiatry, psychology, and psychiatric social work, three of the fields of application of the psychological sciences necessary to carry out the clinical responsibilities of any large neuropsychiatric program. Individuals receiving this inclusive education would be qualified to perform all, or nearly all, the duties involved in caring for psychiatric patients. This is a long list of functions, including arranging for intake, determining the chief complaint, obtaining the present history, taking a physical and psychiatric anamnesis and review by systems, getting corroborative history and other facts from the patient's family and friends or from institutions with which he has had contact; performing physical and neurological examinations; determining the mental status; requesting or performing indicated medical laboratory tests; carrying out necessary psychological examinations; making a diagnosis and prognosis; conducting therapy; directing occupational, recreational, physical, and other adjunctive therapies; doing research directed toward improving available clinical techniques; arranging for disposition of the patient; doing case work with his family and friends as required; and following the patient's later course for purposes of further treatment, maintaining records, and research.

The fact is, however, that it is not feasible for all this to be done by one individual. Even if one person could perform all these tasks at the highest level of professional accomplishment which has been reached—and no one could—it would still be true that division of labor is more efficient, for it has often been shown that a professional team can deal with a larger number of patients than they could if each person worked independently. Specialization has become a recognized necessity throughout the medical sciences, and the complexity of the clinical services which should be offered in the psychological field indicates clearly that there must be specialization here also. For the most satisfactory care of mental patients, the adoption of the neuropsychiatric team, made up of a minimum of three professions—psychiatrists, clinical psychologists, and psychiatric social workers—is inescapable.

Another obvious fact leading to the same conclusion is that there are few individuals of sufficient motivation, economic security, or emotional stability to complete the years of study necessary for such inclusive training. Accomplishing complete education in all these fields is a utopian goal which can be met, at best, by very few.

Let us include specialization in our ideal professional situation, therefore, and attempt to determine the most satisfactory allocation of tasks to each of the three specialist groups. In doing this, we shall neglect such crass human considerations as restrictions of licensing, professional jealousies, and resistance to change, assuming naively that the sole concern of all is the welfare of the patients.

An Effort to Distribute the Duties Ideally. One distribution of responsibility for which precedent could be found would be for the psychiatric social worker to perform necessary intake tasks, make arrangements for disposition of the patient at the end of treatment, and see that he is available for follow-up, leaving the psychiatrists and clinical psychologists to divide responsibility for the other care of the patient. Conceivably, this arrangement might be satisfactory, but it is evident that such a division of labor creates problems.

For one thing, the patient has family and friends and is a component of society. There are reciprocal relations between the patient and those who constitute his environment, and whenever there is mental maladjustment in one person there must also be some maladjustment in those he contacts. If the husband is ill, the wife is affected; if the child has symptoms, they may be the symptoms of his mother's illness. Frequently, arbitrary considerations determine which member of the involved group is considered to be the patient and is treated by the psychiatrist. It may be the one who has the most overt symptoms; the one most willing to consult a doctor; or the one who came to the hospital for a medical condition and was referred to a psychiatrist. It may be that only the veteran in a maladjusted family can be treated, because only a veterans' clinic is nearby; or it may be that only the child can be seen, at a neighboring children's clinic. Traditionally, social workers do "case work," which is often a kind of therapy, with the non-patient members of the family. Since the selection of the patient is so arbitrary, however, no difference should ideally be permitted between the training of persons who treat non-patient members of the family and those who treat the patient. The fact that a social worker has the psychiatrist to turn to when she encounters difficulties may mitigate this inequality, but there is clear injustice unless all members of a maladjusted family group are cared for by individuals with the same sort of training, whatever that may be.

In making disposition of patients, one of the important tasks of social workers in the Veterans Administration and elsewhere is finding them jobs. In an ideal treatment center, persons doing this would have to be conversant with the whole range of possible vocations, know job descriptions accurately, and be able to administer and interpret vocational tests. Social workers are not at present trained in this field and, therefore, in ideal situations vocational psychologists would have to be employed for this part of disposition. For these few reasons, out of many which could be mentioned, we can see that the social worker's job cannot easily be categorized as "intake and disposition."

Recognizing that this is not a satisfactory definition of the scope of the social worker, let us, nevertheless, proceed to consider a possible distribution of the remaining duties, again dealing with a division of labor which can find plenty of precedent. Suppose that, in our clinic, the clinical psychologist does psychological testing and the psychiatrist carries out all other aspects of clinical care, from taking the chief complaint to returning the patient to the social worker for final disposition.

Such an arrangement might be workable if there were not more to psychological diagnostics than administering tests and adding up the scores. Much clinical skill is required in reporting even a simple I.Q. to make allowances for the patient's linguistic handicaps, the extensiveness of his education, his eyesight, his socio-economic background, his physical health, his *rapport* with the tester, and many other considerations. Interpreting the Rorschach test or other projective procedures is much more complex. Relating objective diagnostic findings to the circumstances of the individual case can be done only if the interpreter has had enough clinical experience to understand thoroughly the problems of human personality and mental disease. Interpretation of psychological diagnostic procedures can best be done by the one who has administered them, but if the psychologist has done only routine testing he will be unable to make the most of his procedures.

If the psychologist is to be an ancillary technician in the complexities of psychiatric diagnosis, it must be in the way that the radiologist, rather than the laboratory biochemist, is a technician. In general medicine, it is clearly understood that the man who interprets x-rays must have a broad background including the whole range of clinical experience. The same is true of the psychological diagnostician of the future—the nuances of these diagnostic problems are too many and complex to be learned in the fastnesses apart from patients which can more reasonably be occupied by biochemical technicians. The psychological diagnostician who does not constantly submit his findings to the validation of the therapeutic course rapidly becomes removed from clinical reality, sterile, and esoteric.

Another consideration which relates to the advisability of limiting the role of clinical psychologists to testing and delegating nearly all other responsibilities to psychiatrists is that most training for psychiatrists offers little opportunity for them to become conversant with psychological principles. They may take a course or two in psychology in college and one of the very feeble courses in this subject which some medical schools offer, but they are not thoroughly grounded in the wide field of experimentation and investigation which, at present, forms the systematic science of psychology. The average well-trained clinical psychologist knows more about the "physiology of the mind," the normal functioning of the personality, and related topics than does the average well-trained psychiatrist. At present, psychologists can contribute a great deal to interpretation of all sorts of behavior from such a background, and if the psychiatrists do not have this knowledge available to them, either from their own intensive study or from the comments of colleagues in the neuropsychiatric team, their handling of patients will be less effective. The systematic body of facts collected by investigation in the fields of experimental, physiological, comparative, child, social, and personological psychology are capable of innumerable sorts of clinical application. A beginning has scarcely been made in employing in the clinic the findings of the laboratory on such matters as maturation, perception, learning, memory, motivation, group behavior, and many others.

Because the region of the mental sciences is so little explored, emphasis on research should be of paramount importance in every psychiatric center. Much diagnostic and therapeutic work in psychiatry is so feeble that tremendous investigative efforts should be made to catch up with the other divisions of medical science. In most of these other divisions, there are many M.D.'s or Ph.D.'s trained in laboratory and clinical investigative science who devote a large part of their energies to advancing the field. Some of these not only understand experimental procedures but also have had clinical experience, which makes them more effective. The vast majority of psychiatrists lack the necessary training for research, and, because experimental institutes and laboratories in psychiatry are almost unheard of, it is understandable that their training has not motivated them for it. Many psychologists—by no means all—on the other hand, because of their professional philosophy and traditions, are intensely eager to do research. They are acquainted, by experience, with a number of methods applicable to clinical investigation concerning which most psychiatrists are ignorant. Unfortunately, they usually have not tried to get, or have not been permitted to obtain, sufficient clinical experience to enable them to apply this knowledge effectively. In their ignorance, they are likely to be supercilious and unreasonable in the demands for precision which they make of clinical research. Unless they take an integral part in clinical operations of all sorts, they will not be able to get the experience which will enable them to make the most satisfactory applications of their method and to fit it into normal clinical routine so that it will be beneficial rather than harmful to individual patients. An alternative is to make psychologists out of psychiatrists, so that they will be equipped with this technical background so valuable for research.

The conclusion to be drawn from this discussion of the ideal arrangement for rendering psychiatric services is that some degree of specialization is desirable from the point of view of efficiency and thoroughness of operation. However, it is not possible or desirable, even in an ideal setup, to delimit sharply the activities of social worker, clinical psychologist, and psychiatrist. It is important for them to have certain types of overlapping functions. Such overlap must be recognized in ideal job descriptions for the three professions. When we can approximate such descriptions, we may proceed to consider what education each group should have in order to fulfil its ideal role. Finally, we may try to discover how closely it is possible, at present, to attain such ideal training.

Differences in the Backgrounds of the Professions. Let us deviate, at this point, to observe the present differences in background among the three professions. First, there are striking differences in education. Psychiatric social workers complete two years of graduate work to a master's degree. Clinical psychologists now usually complete three or four years of graduate education to a doctor's degree. Psychiatrists complete four graduate years for the M.D. and then have one to six years of further training, largely clinical. This is a wide discrepancy in the amount of preparation.

Characteristically, schools of social work base their curricula on the assumption of undergraduate training in related fields. This has the effect of lengthening professional training by beginning it before the graduate years. Most of these schools prefer or require for matriculation some systematic sequence of courses in the social and biological sciences, including sociology, anthropology, psychology, history, government, economics, and biology. The graduate work founded on this includes courses, reading, and field exercises designed to give training in the following areas: psychology, usually so taught as to emphasize the dynamics of personality mechanisms, and commonly not presented at an advanced level or by highly qualified psychologists; public welfare and administration; community organization; social agencies and institutions, public and private; social statistics, including extremely elementary instruction in research techniques and in the sorts of procedures necessary for compiling socio-economic data into tables and charts; elementary medicine, introductory psychiatry, and basic legal concepts relevant to social work; and economics.

The most significant part of the preparation of psychiatric social workers, however, is not courses but carefully supervised practical training in psychiatric case work. This involves detailed analysis of interviews, in many ways like the apprentice training of some psychiatrists, as well as practical direction as to how the many resources at the command of the social worker can be most effectively employed for the welfare of the patient. Through an entire year of supervised internship, the student, under direction, learns by his own mistakes. Throughout all of this, a dynamic approach to human personality is adhered to, which is more commonly than not doctrinaire, being cast into the Freudian, the Rankian, or some other theoretical system. While the graduate training of social workers is officially two years in length, actually it has been traditional in the profession to continue this apprenticeship into the early years of full-time employment. This on-the-job training directed by case supervisors, even if not of formal character, is nevertheless an important educational experience, and it is important to realize that the two graduate years in social work are by no means the end of the preparation of a conscientious worker.

A final item in the education of a social worker often has been a psychoanalysis. Some schools have strongly opposed such training, while others have exerted such pressures in favor of it that their students have felt it to be essential, although there has not been any official requirement. Commonly, such pressure has been for personal rather than didactic analyses, and most of the analyses of social workers have not been of the training type, though of course they could not help but have educational importance. Though there has been, in recent years, perhaps more general acceptance of the value of psychoanalyses, it is probable that a smaller percentage of social workers are receiving them now than were ten years ago, because they are at present so difficult to obtain.

The training of clinical psychologists involves a markedly different subject-matter. Specific course work usually begins in undergraduate college and continues into graduate school. It includes beginning, abnormal, and experimental psychology; psychology of learning, perception, motivation, and the higher mental processes; theoretical systems in the psychology of personality and clinical psychology; mental hygiene; the administration of objective tests of intelligence, attitudes, aptitudes, and other traits; the use of projective techniques and other diagnostic procedures including mental status examinations; theory of interviewing and psychotherapy; statistics; theory of test construction, experimental design, and scientific method. Practical experience is given in the construction of educational, vocational, industrial and clinical testing procedures; in the use of all these methods; and in the conducting of independent research. Supervised internships of various lengths usually involve the application of diagnostic rather than therapeutic methods. Frequently, courses in related fields such as human biology, physiology, cultural anthropology, and sociology are required. There is a highly academic insistence on developing skills in foreign language and on preparing a dissertation. Much greater emphasis is laid on preliminary, comprehensive, and oral examinations than in the schools of social work or psychiatry. Occasionally, clinical psychologists consider a psychoanalysis important in their education, but this is not emphasized as much as it is in social work or psychiatry.

The first observation we make in considering the training of psychiatrists is that there are several accepted routes to becoming a member of this profession. All have the first portion in common, which is the four years of medical school. In typical medical schools during these four years, usually less than one-twentieth of the time is devoted to clinical psychiatry, and occasionally a brief course in psychology is added. Certain skills and attitudes gained in practicing clinical medicine can be transferred to the practice of psychiatry and are of great importance. These include knowledge of the nature of illness; experience as to how human beings of various types react to their diseases; a comprehension of how to approach patients; an understanding of clinical method; an ability to bear responsibility for decision concerning their welfare—a sort of *esprit de corps* of medical integrity comparable to the morale inculcated into the cadets at West Point. Though they do obtain much relevant to psychiatric practice in medical school, it is a fact that, at present, medical students learn little direct psychiatry beyond an elementary understanding of diagnostic categories.

Three common roads lead to a career in psychiatry after medical school. The first includes a medical, neurological, or rotating internship together with further training in neurology. The primary interest of such doctors is in neurology and neuropathology, but, since they find that a large proportion of their practice is psychiatric, they are forced by practical considerations to call themselves neuropsychiatrists. Having had little psychiatric instruction, they usually take a "common sense" approach to such problems and rely heavily on advice, reassurance, vitamins, and trips to the country. With their "organic" neurological background they frequently express outspoken antipathy for any other type of psychiatry.

Second, there is the "state hospital" approach to psychiatry. After medical, rotating, or other types of internships, these physicians go to hospitals where most of the patients are psychotic. Their interest becomes largely focused on diagnosis, often in Kraepelinian terms, on legal questions, custodial care, and final disposal of the patient. It is the sober truth that, with the exception of shock treatment and a few other less widely employed procedures like prefrontal lobotomy, there is and has been no serious therapeutic attempt in most of these hospitals, largely because of the unresponsiveness of these illnesses to therapy and because the doctors have been so overworked in their other duties.

Third, there is the "dynamic" approach to psychiatry. This postgraduate training usually consists of a didactic analysis and then an apprenticeship under an analytically-minded psychiatrist—review of interviews, suggestions on how to conduct the psychotherapeutic course, and "control analyses." In a few rare residencies only can the student get a broad selection of opinions and theories, including more than one man's approach to psychotherapy or the doctrines of more than one school of psychiatry. Psychoanalytic institutes, however, conduct courses and seminars in which the viewpoints of a number of instructors can be obtained, but the range of their attitudes is usually limited by their acceptance of a common basic doctrine. Rare is the inclusion of psychology, sociology, and anthropology in graduate psychiatric training.

Not only do the three professions of psychiatry, psychology, and psychiatric social work at present vary widely in their training, but also, largely because of these training differences, they have markedly divergent motivations. A point of importance about social workers is that most of them are women. This largely affects the professional motivation of the group, just as it does the attitudes of the nursing profession. Because they are women, they accept more easily working as submissive subordinates to the doctor and being constantly under the medical aegis. This makes for good social relations in the social worker-doctor team, but is not unequivocally good for the profession. When, during the war, the percentage of males acting as social workers became large in the Armed Forces, there was a marked increase in friction between the two professions. Physicians from time to time said they were distressed at the "new pretensions" of social workers, and the social workers for the first time pressed for a clear delineation of their duties, which before had not seemed of primary importance. Since most psychologists are male, it is likely that there will be more competition between them and psychiatrists than there has been in the past between social workers and psychiatrists.

A motivational undercurrent of social work, which springs partly from the fact that a high percentage of the profession is feminine and partly from the fact that the profession is constantly in contact with official and semi-official social agencies, is a strong tendency to work toward maintaining the status quo of the social order. An important responsibility of social workers is to help order the environment in which the patient will live, and make disposition of the patient at the end of treatment in the way best calculated to do away with conflict. An important aspect of this is arranging it so that the patient lives in harmony with his society, and to do this the social worker is motivated to continue the recognized patterns of behavior in societies, whether or not he may intellectually wish to accomplish reform by community education or legislation. This is a markedly different approach from that of most psychiatrists or psychologists, who operate in private relationships which do not exert pressures toward conformity as the public contacts of the social worker do.

The largely masculine motivation of psychologists differs vastly from that of social workers. Among clinical psychologists, there is undoubtedly a greater idealistic urge to help suffering individuals than among any other group of psychologists, but even this idealism is commonly expressed by them in terms of the discovery of generalizations about human personality and mental disease which later is capable of application in the individual case. Psychology has dwelt most of its life amid academic towers. It has long professed an idolatry of exact science, of predictable law, and of research to achieve these ends. The main stem of psychology has by no means felt itself to be a step-child of medicine, nor has it envied the role of the physician, except for his prestige in recent years. Rather, older psychologists have viewed clinicians haughtily as being beneath them because they are inexact impressionists. They have yearned to approach the precision of physiology, hoping some day to scale to the heights of accuracy of organic chemistry, physics, astronomy, and, above all, pure mathematics.

Even a superficial survey of the history of psychology reveals important facts about present motivation of psychologists. The metaphysical beginnings of psychology as a contemplative discipline were soon overcome by the development of a laboratory science patterned after physiology, begun by Wundt and Helmholtz in the middle of the last century. Until quite recently, this experimental approach has constituted the central stream of psychology. Emphasis has been placed on psychophysics, sensation and perception of external stimuli, learning, memory, attention, feeling, animal behavior, and similar fields in which precision of measurement has been possible. Studies of abnormal behavior, like those by Charcot, Prince, Janet, and Freud, have always contributed to psychological thought, but these have not received the main emphasis. Similarly, interest in group phenomena and the applications of psychological principles to masses in industrial and personnel psychology has been a sideline. In the last decade, there has been a tremendous increase in concern with psychoanalysis and all its offshoots, with the counseling techniques represented by Carl Rogers and others, and with the personalities and mental illnesses of individuals, so that during the war years there has been a real revolution, and the clinical psychologists, who were once underdogs because they dealt in fields where precision was difficult, have now become "top dogs," although the demand for precision and careful method still is loud and there are many different points of view in psychology. This divergence in viewpoint among psychologists is diluted by the extremely broad range of matters with which they concern themselves, a scope far greater than the treatment of ill or socially maladjusted individuals, which is the primary concern of psychiatrists and social workers. Today, psychologists are writing advertisements, choosing personnel, counseling normal individuals on vocational and marital matters, diagnosing and treating mentally ill patients, studying characteristics of human sensations, investigating public opinion, gauging social trends, determining principles of proper education, recommending to engineers how to arrange instrument panels, designing houses to suit human limitations and sentiments, analyzing administrative organizations, and devising methods for preventing fatigue and safety hazards, as well as carrying out many other diversified activities.

The professional purpose of psychologists, then, clearly differs from the aims of the other professions, though the recent revolution has brought it closer rather than farther away. It differs from the well-defined goal of the medical psychiatrist and social worker, which is to help the human being who needs aid now. The psychologist characteristically hopes to determine principles of normal and abnormal behavior, with emphasis on the normal, as precisely as possible, with the expectation that once these laws have been determined they can be applied for the welfare of men.

Distribution of Duties Forced by Present Circumstances. We can now consider the training situation in clinical psychology. What sort of education in this field is it necessary and feasible to ask a university to give at the present time?

Let us hope, that, by now, most trade-unionism can be dispersed; that professional antagonisms will not be so bitter as to be determining factors; that legal restrictions can be surmounted or circumvented; and that assignment of diagnosis, therapy, research, and other clinical activities can be to the best qualified person, whoever he may be. Let us, therefore, neglect such possible hindrances in our present considerations.

It is impossible, however, to sidestep the tremendous demand which exists and which will continue to exist indefinitely for services in clinical psychology. The requirements of the Armed Services in the last war found the country to be shorter of psychiatric service than of any other medical specialty. If the country is to continue adequate mental hygiene and hospitalization programs in psychiatry, the supply of psychiatrists will be far short for many years. Group therapy alone does not give promise of settling the problems of how psychotherapy will be given to the masses who need it. Inevitably, whether we want it or not, psychiatric social workers and clinical psychologists will be called upon to assume therapeutic roles.

A report on an official Army survey of duties performed by clinical psychologists in the Army during the last war indicates that, from January 1945 to April 1946, about one quarter of the time of psychologists was spent in doing guidance and therapy. The authors of the report state:*

"Most of the time devoted [by clinical psychologists] to therapy was for individual therapy; the rest of the time was given to counseling and to group psychotherapy. It is significant that these therapeutic efforts were carried on under the direct supervision of the neuropsychiatrist. While, in part, this function represented the desires of psychologists to participate in this work, in much larger measure it was the result of a tremendous patient load in neuropsychiatric and in neurological sections which could not be handled by neuropsychiatrists alone, and so involved both clinical psychologists and psychiatric social workers. As psychologists gained experience along this line, they were used more frequently to meet the emergency needs. The functions of re-education carried by psychologists with aphasic patients and with other physically handicapped patients (such as the paraplegics and the deafened) contributed to the relatively large amount of time devoted to guidance and therapy.

---

* Hutt, M. L., & E. O. Milton. The duties performed by clinical psychologists in Army Medical installations. Bull. Mil. Clin. Psychol. 1: 115, 117. 1946.

"Many of the narrative comments supplied by psychologists indicated that they were more interested in careful diagnostic work and particularly in clinical research in psychodynamics and psychotherapy than in therapy as such. In these areas . . . the psychologist, basing his service on solid training and scientific method, has a far more important contribution to make [than in psychotherapy], especially under less pressured situations."

These last statements undoubtedly represent the majority opinion of qualified clinical psychologists. In general, they believe that their most important contributions can be in fields other than direct therapy, and therefore they hope to evade extensive involvement in psychotherapy. It must be admitted that there are a number of clinical psychologists whose chief purpose in entering the field, like that of some psychiatric social workers, is in order to do therapy—but this is a minority, and the developmental history of psychology makes it clear why this is so. Anyone following non-medical roads to therapy merely because they are shortcuts (though clinical psychology is not a very short shortcut) of course deserves disapproval.

Is it not reasonable to ask here, however, whether a graduate clinical psychologist or psychiatric social worker is not better trained to do "dynamic" psychotherapy than the "psychiatrist" who has chiefly a neurological background, the "psychiatrist" with a custodial, state hospital background, or the general practitioner? Some will argue that the skills transferred from general medicine to such clinical work are so important that any physician is better trained for such therapy than any layman. One bit of evidence advanced to support this position is that the young medical officers in the Army just out of their internships who were given ninety days of psychiatric training were thought by some to make, in general, better therapists than experienced clinical psychologists. If this judgment is correct, it still would be difficult to tell whether this is the result of transfer of general medical training or the consequence of the fact that the prestige of medicine in our society and the excellence of medical aptitude tests together serve to select for medical training young men of outstanding all-round ability who therefore are capable of superior performance in any task they undertake.

The demand for psychiatric treatment is so great that ten times as many psychiatrists as we now have probably could not handle it adequately. Therefore, the doctor must decide which parts of the therapy in this field he wishes to handle himself. Does he wish to do occupational therapy, recreational therapy, habit retraining, re-education of aphasics, other speech therapy, counseling on normal adjustment problems, or vocational counseling? Almost all of these therapies are already recognized to be the province of psychologists or other technicians, and very little complaint is heard from physicians. There can be no doubt that these are therapies in a respectable sense of the word.

Even when the psychiatrist delegates all these functions, he still has an enormous task in the direct treatment of the more severe mental maladjustments. The jokes about patients growing long white beards waiting to be analyzed and the comments about only the rich being able to afford neuroses are not humor untouched with reality. If America is going to make a serious effort in the next decade to give therapy for mental maladjustments where it is needed, large numbers of psychotherapists must be trained—more than medical schools with their present limited enrollments will be able to turn out. There are three possible sources: multiply medical school enrollment at least by two; train lay psychotherapists; or evade the responsibility to give help where needed, as, for instance, we have neglected the scourge of pellagra in certain sections of the South. The first course seems highly improbable, at the moment, and the last would have many serious consequences, not the least of which would be the turning of the ill to the psychological cultists.

Even if this problem were not so critical, the reasons mentioned above as to why a psychologist should do therapy to improve his diagnostic and research skills are cogent enough. The conclusion, therefore, must be that for both idealistic and realistic reasons the ideal university will train clinical psychologists in coming years to carry out the following three functions: diagnosis, research, and therapy.

A Specific Training Program. In planning education in the various psychological professions, we must accept the principle that each of the three chief professions involved should have different training, which follows directly from our earlier decision that there should be specialization.

We shall not go into detail as to what should be required for social workers, except to say that, if they are to do case work such as they have done in the past, which often means carrying out extensive therapy, they should receive longer and more intensive formal training than they have in the past. This preparation should be based on the work they are to do, and if their responsibilities are to include therapy, they should have the necessary broad background for this, as well as supervised practice (in which social work has been stronger than either psychology or psychiatry). It is not fair to the patient to have him treated by unskilled therapists, using as a rationalization the semantic distinction between case work and therapy.

In the training of psychologists and psychiatrists, there should be as much merging of backgrounds as possible. It would be well for modern universities to develop graduate schools with curricula permitting a number of individuals who wish to have really thorough preparation to become qualified in both psychiatry and clinical psychology. A possible schedule might be something like the following: two years of liberal arts college; one year of advanced clinical psychology, sociology, and cultural anthropology; one year of the pre-clinical medical sciences which would be comparable to the present first year of medical school. At the end of these four years, a bachelor's degree would be granted, following which the candidate would take the second and third years of medical school work and then spend a year in psychiatric clinical work in a general hospital, in a mental hygiene clinic, or in a neuropsychiatric hospital. At the end of this time, he would be granted the M.D. degree in psychological sciences. He then would have one year of a rotating psychological-psychiatric internship, which would include both psychological examinations of all sorts and the performance of various psychiatric duties, complementary functions of equal status. This would be followed by one year of independent research leading to a dissertation, seminars, and perhaps a psychoanalysis. At the end of this year, a doctoral degree in clinical psychology might well be awarded. Work from this time until the candidate qualified for his specialty boards in psychiatry or clinical psychology might consist of mixed psychological-psychiatric residencies, including work with psychoneurotics, psychotics, psychosomatic patients, and clinical research.

There would necessarily be other individuals, and in larger numbers, who would want either a psychiatric education alone or a psychological education alone. The same graduate school that set up the combined program outlined above should be able to give both these latter courses of training, because even if the curricula are separate they should be closely integrated, since clinical psychologists and psychiatrists should learn from the beginning of their training to operate as a team.

The report on graduate training in clinical psychology made by the Subcommittee on Graduate Internship Training of the American Psychological Association and the American Association of Applied Psychology outlines the most satisfactory plan yet suggested.* It begins by suggesting that the undergraduate program of students in this field should include 20 semester hours in psychology, to consist essentially of the standard courses for undergraduates, not including professional or advanced work—courses like beginning psychology, beginning abnormal psychology, psychological statistics, learning and perception, the higher mental processes, physiological psychology, etc.; 20 semester hours in biology, physics, and chemistry; 9 semester hours in mathematics, through introductory calculus and statistics; 9 semester hours in the fundamentals of educational philosophy and experimental didactics; 12 semester hours in anthropology, sociology, economics, and political science; 6 semester hours in the history of culture, philosophy, logic, and comparative literature; and enough work in modern languages to develop a reading knowledge of two, preferably French and German.

---

* Subcommittee on Graduate Internship Training to the Committee on Graduate and Professional Training of the American Psychological Association and the American Association for Applied Psychology (D. SHAKOW, Chairman). J. Consult. Psychol. 9: 243-266. 1945.

Building on this general background, the report suggests a four-year graduate program leading to the doctorate, including a full year of internship, preferably the third year. The courses recommended for the first year are general psychology, 6 semester hours; dynamic psychology I, 3 semester hours; experimental clinical and dynamic psychology, 3 semester hours; developmental psychology, 6 semester hours; theory and practice of psychological tests and measurements I, 3 semester hours; physiological sciences, including physiology, anatomy, neuroanatomy, neurophysiology, and endocrinology, 6 semester hours; advanced statistics and qualitative methods, 3 semester hours. In the second year, the committee recommends dynamic psychology II, 3 semester hours; experimental clinical and dynamic psychology II, 3 semester hours; theory and practice of psychological tests and measurements, 6 semester hours; theory and practice of projective devices, 3 semester hours; therapeutic theory and methods, 6 semester hours; methods of case study and analysis, 3 semester hours; introduction to clinical medicine, 3 semester hours; educational and vocational guidance techniques, 3 semester hours. The third year consists of an internship with various sorts of psychiatric patients. The recommendations for the fourth year include independent research leading to a dissertation; cross-discipline seminars attended by representatives of psychology, anthropology, sociology, psychiatry, etc.; seminars in professional problems, standards, ethics, etc., of psychology; and additional courses in psychology as needed to round out the individual's curriculum.

The new thinking contained in this report is evidenced in various ways: by inclusion of pre-clinical sciences; by emphasis upon training in medicine, neuroanatomy, and neurophysiology; by emphasis upon related fields like sociology and anthropology; and above all by stressing the clinical rather than the academic approach, following medicine in teaching as much as possible at the patient's bedside. It would not, however, constitute adequate preparation for independent psychotherapy, which would have to be obtained later if required.

The most satisfactory way to give such a psychological curriculum might well be in a separate graduate school for the psychological sciences. Here, certain basic graduate courses could be given, and then specialization could be undertaken in medical or clinical psychology, in pure research and experimental psychology, in social psychology, in industrial psychology, or in other areas. There should be close liaison with the medical school, the law school and the business school, and combined programs with all of them.

Psychiatric education needs as radical alterations of the traditional program as does psychological education. The changes should begin in undergraduate medical school. First, an inclusive and well-planned course in psychology, including laboratory work, should be presented in the first year in the same status as the other pre-clinical sciences of normal function, like biochemistry and physiology. Second, there should also be, to a lesser degree, opportunity for study in sociology and economics. Third, since at least half the cases seen by a general practitioner have significant psychiatric aspects, certainly psychiatry should be taught in every one of the clinical years and should include much more emphasis on psychoneurosis and psychosomatic medicine than on psychosis. Fourth, there should be a real course in applied clinical psychology, including psychometric testing, the use of projective techniques, and all other methods available to clinical psychologists. Medical students should have an opportunity to use these procedures in order to familiarize themselves with the theory behind them and to understand the significance of reports based on their use. Fifth, there should be courses in biostatistics and in the scientific method. Sixth, medical students should be required to do independent investigation, resulting in a dissertation in a field of their choice, so that they will have respect for, motivation for, and understanding of the importance of research. Last, medical students should have the opportunity to work as members of the neuropsychiatric team.

After medical school, interns and residents in psychiatry should learn more than how to make diagnoses and give shock treatment. There should be instruction in psychoanalytic and other dynamic psychiatric concepts. A valuable adjunct to this is the study of literature, art, and history from this dynamic viewpoint. There should be instruction in psychotherapy both in an apprenticeship under an individual instructor, and in groups making detailed studies of the therapeutic progress of individual cases under the direction of various psychiatrists, so that different approaches can be learned. Student psychiatrists should become proficient in organizing and participating in the neuropsychiatric team. Finally, emphasis upon clinical research should be markedly increased in psychiatry, because of the great need for advancement in this field.

It is clear that all the professions in the psychological field will in the future be working together. Thousands need to be trained in each of them. The educational programs must be united in a pattern as closely integrated as the cooperative activities in which their graduates will take part.